Origin, Composition, and Role of Antimicrobial Plant Resins Collected by Honey Bees, *Apis mellifera*

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Michael Bradley Wilson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser: Dr. Jerry D Cohen

May, 2014

Acknowledgements

I would like to thank everyone who has supported me through my sometimes bewildering six and one-half year journey through graduate school – From toxicology to plant sciences. It has been a privilege to work with all the bright minds I have encountered in the Department of Horticultural Science, the Department of Entomology, and the Plant Biological Sciences graduate program. I would like to give a special thanks to the past and current members of the Cohen, Gardner, Hegeman, and Spivak lab groups. Specifically, I would like to thank Doug Brinkman and Dr. Alison Pawlus for their contributions to my work, as it would have been a much longer road without their help.

I want to thank my advisor, Dr. Jerry Cohen, for giving me the opportunity to join the Plant Biological Sciences graduate program. I was extremely close to cutting my losses and leaving the Twin Cities when I was on the ropes in toxicology, and I don't know where I would be if he had not recruited me. To the other faculty members of my Ph.D. committee (Drs. Florence Gleason, Gary Gardner, Adrian Hegeman, Marla Spivak) – Thank you for being so involved in my research and showing such an enthusiastic interest in my work. I cannot imagine a more willingly collaborative and interactive group of faculty, or a graduate student that has worked more closely with all of their committee members.

There is not enough paper in the world for me to properly thank my friends and family for all their love and support, but I must try. Simply put, none of this would have been possible without my parents, Brad and Laurie Wilson. They worked hard to allow me every opportunity to make something of myself. Through them I learned (in very different ways with respect to each) toughness, perseverance, determination, effort, self-discipline, humility, kindness, and, most importantly, to give respect freely instead of foolishly making people try to earn it. Teaching me to be a good person was always their highest priority, but I know by mom is secretly very proud to have raised a doctor - even if he turned out to be a Doctor of Philosophy! And where would I be without my brother and sister, Nicholas and Amber? Lonely and bored, that's where. To them I am just their slightly nutty older brother – for which I am eternally grateful. I have too many family

members and too little space to thank everyone properly, but I would also like to give special thanks to my Grandma Cunconan and Grandma Wilson – I hope I make you proud. With married life comes an entirely new family, and I am happy to thank the Pardo clan for copious amounts of Brazilian food to fuel my endeavors and their absolute acceptance of me into their family.

I have more good friends than any person should be rightfully blessed with, and I tip my largest hat to Mike, Jason, Kyle, and Jeremy (and their families) with whom I have been friends with for the vast majority of my life. Together we learned how to fight, fly, and crow and I could not imagine this adventure without them. Thank you also to my friends Kirk, Clint, and Walker from Simpson College and beyond, who have helped me through this strange and wonderful journey of life. I thank my high school mentors John Elkin, Nancy Lehman, and John Crippen for making it impossible not to like school, and I also thank my undergraduate mentor, Dr. Ron Warnet. Without his encouragement I would not have gone to graduate school. Special thanks to Master Noble Ram and the people of Ram's Taekwondo for kicking me hard enough to make me forget about graduate school for a little bit. I would also like to thank my tiny dog and official writing buddy Peppy, whose head lay in my lab while writing the majority of this dissertation. His silent encouragement was ever a relief for writer's block.

Lastly I would like to thank my beautiful wife Fern (Yes, like the plant) for putting up with a slightly unconventional and quirky Ph.D. student for the past three years. Without a doubt, the best thing I have done in the Twin Cities is meet her. She makes me want to come home from the lab to see her, and I don't think there is any higher complement that can come from a scientist. I look forward to being your slightly unconventional and quirky *something*, whatever that may be, for many years to come.

Again, thanks to everyone!

Cheers,

Mike

Dedication

This dissertation is dedicated to all the people who have supported me – but particularly to my grandfathers, Edward Cunconan and Alfred Wilson, and my cousin, Bobby Karpan, who could not see its completion.

Abstract

The use of plant resins in nest building by honey bees, *Apis mellifera*, is an oft ignored, but critical behavior to bee health. Termed 'propolis' by beekeepers, deposited resins in the nest have many positive physiological effects on the colony. Honey bees are feral and abundantly managed in many regions of the United States, from the Sonoran Desert to icy Minnesota, yet a diversity of very different resinous plants exist in every environment they call home. We know very little about what resinous plants bees utilize in these different regions, or what benefits bees might derive from specific plants. It is thought that the antimicrobial properties inherent in resins, which are complex mixtures of phenolic and isoprenoid compounds, are important drivers of their derived benefits to bees. The research herein focuses on creating better methods to track resin forager behavior, and then using those methods to discover the botanical sources of bee-foraged resins, while also exploring how resins from different plants directly affect the growth of two bee pathogens, the gram-positive bacteria *Paenibacillus larvae* and the fungus *Ascophaera apis*. I found that individual resin foragers can be chemically tracked to their resinous plant targets using metabolomic methods that hold great advantages over traditional chemical analyses, and that there is much diversity in the ability of resins from different *Populus spp.* to inhibit the *in vitro* growth of *P. larvae* and *A. apis*. I go on to further explore the benefits of different resins and find that propolis from Fallon, NV was particularly active against *P. larvae* and *A. apis* out of samples from 12 different regions in the U.S. Finally, I used bioassay-guided fractionation against *P. larvae* to isolated several flavanone-3-alkyl esters from NV propolis that displayed very high activity (IC_{50} = 17 µM to 68 µM) against *P. larvae* and *A. apis*. Re-examination of data from my previous studies indicated that these compounds were strong contributors to overall *anti-P. larvae* activity in regional propolis samples, and that *Populus spp.* are likely the botanical sources of these compounds.

Table of Contents

List of Tables

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 1

An Introduction to Plant Resins and Honey Bees

The importance of bees to plant-based agriculture and their current status

Honey bees, *Apis mellifera*, are the world's most important pollinator of agricultural crops (Klein et al., 2007). While honey bees are not required for the production of wind-pollinated staple crops like wheat, maize, or rice, they are required for the production of fruits, nuts, forage, and vegetables. The U.S. National Research Council's Committee on the Status of Pollinators in North America places the value of honey bee pollination between $15 – $20 billon dollars per year in the U.S., with conservative estimates of indirect value (e.g. the value of seed crop-dependent industries) (National Research Council, 2007). Production data from 1996-1998 indicate that honey bees contribute most to the value of alfalfa ($4.65 billion/yr), apples ($1.35 billion/yr), almonds ($960 million/yr), cotton ($857 million/yr), citrus ($834 million/yr), and soybeans ($824 million/yr) in the U.S. (Morse & Calderone, 2000; Johnson, 2010). Recalculation of these values using current production would almost certainly lead to a much larger overall estimate, mostly driven by recent growth in the almond industry.

Crops vary greatly in their dependence on honey bee pollination, though exact dependences as a function of production value remain somewhat uncertain (Allsopp et al., 2008; Gallai et al., 2009), likely because pollination dependency varies by crop cultivar and growing region. For example, the value dependency of citrus on insect pollination is between 20-80% and the proportion of that pollination done by honey bees is between 10-90% in the U.S. (Morse & Calderone, 2000; Johnson, 2010), making it difficult to calculate a nationwide value dependency. Another complicating factor is the indirect value that can be attributed to pollination services, as several other agricultural industries are largely dependent on honey bee pollinated seed crops. For example, the beef and dairy cattle industries in the U.S. are dependent on alfalfa seed production, which is made possible by honey bee pollination, and the value of beef and dairy potentially add a very large amount to the total value of honey bee pollination services.

The considerable economic value of honey bee pollination makes the on-going decline in commercial pollination availability a huge problem for U.S. agriculture, with registered colonies having decreased by 61% in the U.S. from 1947 to 2008 (vanEngelsdorp & Meixner, 2010). The recent, and perhaps most devastating, declines in honey bee colonies have been mostly due to unsustainable winter losses (vanEngelsdorp & Meixner, 2010; vanEngelsdorp et al., 2012). Beekeepers surveyed by the Bee Informed Partnership (www.beeinformed.org), a collaborative group of universities and the U.S. Department of Agriculture, reported average nationwide colony losses of 30.5% from 2006 to 2013 (unpublished data, beeinformed.org). Beekeepers reported total percent colony losses (all reported colonies in a given state) of 70% in Wisconsin, Iowa, and Georgia, while reported total losses in the state of Minnesota were 40% in 2010-2011 (vanEngelsdorp et al., 2012). This is truly stunning, as the approximate acceptable winter loss for a commercial beekeeper is typically considered to be only 15% (www.beeinformed.org).

Almond production is the rightful 'poster-child' for highlighting the economic threat posed by bee decline. California currently produces 80% of the world's almonds, and bearing acreage has exploded in the last decade from 550,000 acres in 2003 to 810,000 acres in 2013 (USDA-NASS, 2013a). During this same period, the production value of almonds has almost tripled from $1.6 billion in 2003 to $4.3 billion in 2012 (USDA-NASS, 2013b). Increased production has led to increased demand for pollination services, as almonds must be insect pollinated to bear fruit, and 90% of this pollination is done by managed honey bees (Morse & Calderone, 2000). Every spring, more than half of all U.S. colonies are trucked great distances to Californian orchards in February to pollinate the blooming almond trees (Sumner & Boriss, 2006). This increase in demand for pollination services, coupled with continued declines in the number of colonies available for pollination, has caused pollination fees to skyrocket from $108 per acre in 2003 to almost $400 per acre in 2012 (assuming two eight-frame colonies per acre) (Sumner & Boriss, 2006; Oliver, 2012). Pollination fees for most crops have recently increased overall, but crops in competition with almonds for pollination services, such as plums, cherries, and avocados, have seen significantly greater increases (Sumner &

Boriss, 2006). Increased costs for growers means increased costs for consumers, and these price increases could further restrict the economic availability of fruits, nuts, and vegetables for vulnerable consumers.

It is important to briefly discuss the importance and status of native U.S. bees as a critical background to put in context the work reported in this dissertation. Native bees are also important agricultural pollinators that are in decline, though we know much less about the extent or severity of their decline (Goulson et al., 2008; Potts et al., 2010). It is estimated that native bees provide $3 billion yearly in agricultural pollination services, although intense, large scale mono-culturing of crops severely limits the pollination impact of unmanaged bees (Kremen et al., 2002; Losey & Vaughan, 2006; Ricketts et al., 2008). Many native bees are highly co-evolved with specific crops and are much more efficient pollinators of their respective plants compared to honey bees on a per bee basis. Specific examples of this relationship between native bees and specific crops include both squash bees (*Peponapis spp.*) with cucurbits and leaf-cutter bees (*Megachile rotundata*) with alfalfa (Canto-aguilar & Parra-Tabla, 2000; Cane, 2002). Another example is the buzz-pollination of field-grown tomatoes by bumble bees (*Bombus spp.*) which can increase fruit set by 45% while also doubling fruit weight, a pollination task honey bees cannot perform due to the morphology of tomato flowers (Greenleaf & Kremen, 2006). Unlike honey bees, native bees will pollinate during cool or rainy weather, which is important to maximize pollination if there is non-ideal weather during blooms. Since the lack of suitable nesting habitat due to farming practices is often limiting to crop pollination by native bees (Kremen et al., 2002; Shuler et al., 2005; Ricketts et al., 2008), the future of cost-effective pollination services may lie in the integration of native bee habitat on or near cropland and the adoption of other bee-friendly farming practices.

Honey bee health

In what seems like an economic climate supportive of commercial beekeeping for pollination services, colony losses due to poor honey bee health remain a systematic and chronic barrier for the beekeeping industry. The importation of live bees into the U.S.

has been prohibited since 1922 (US Code, Title 7, Part 322), yet this ban has markedly failed in keeping dangerous honey bee diseases and pests out of U.S. colonies. The USDA funded Managed Pollination Coordinated Agriculture Project reported 60-100% mortality of stationary colonies not treated for pests or disease from 2009-2011 among apiaries in Maine, Washington, Pennsylvania, Texas, Florida, and California (Drummond et al., 2013). This study displays the extreme dependency of managed bees on human intervention in the U.S. Honey bees and their colonies are parasitized by at least 31 diverse organisms including a variety of bacteria, fungi, protozoa, viruses, mites, and other insects (Morse & Flottum, 1997; Evans & Schwarz, 2011). Although few pests or pathogens actually threaten beekeeping in their own right, the majority of honey bee pests and pathogens only become a significant problem in otherwise weakened colonies.

It is becoming clear that no single factor is responsible for recent large-scale honey bee losses, but interactions between and among factors such as poor nutrition, pesticide exposure, and disease are the likely culprits (Cox-Foster et al., 2007; Alaux et al., 2010; Evans & Schwarz, 2011; Di Prisco et al., 2013; Johnson et al., 2013; Pettis et al., 2013). For example, sub-lethal agricultural pesticide exposure increases viral replication in honey bees (Di Prisco et al., 2013) and also increases susceptibility to the gut pathogen *Nosema ceranae* (Pettis et al., 2013), while both viruses and *N. ceranae* have themselves been linked to large-scale colony losses (Cox-Foster et al., 2007; Higes et al., 2008). Honey bees are intentionally and unintentionally exposed to many pesticides, and it is now known that miticides applied by beekeepers to control honey bee pests can cause up to a 100-fold increase in the toxicity of fungicides applied by growers in the field (Johnson et al., 2013). The factors affecting honey bee health are very complex and systematically difficult to address due to the number and diversity of stakeholders involved, from commercial beekeepers to growers, pesticide manufacturers, and government agencies.

Despite the complexity of the problems involved in honey bee health, a better understanding of the natural mechanisms by which honey bees resist disease and then, in turn, leveraging these traits in managed colonies could lead to strategies that increase overall bee health. While the honey bee genome only contains about 1/3 of the total

genes known to contribute to individual immune function compared to those found in the solitary insects *Drosophila* and *Anopheles* (Evans et al., 2006), honey bees have a sophisticated group of cooperative behavioral traits that contribute to disease resistance on the colony level, termed 'social immunity' (Cremer et al., 2007). An example of a social immunity trait is hygienic behavior, where honey bees work together to detect and remove diseased brood from the nest, resulting in colony-level resistance to a variety of diseases and pests including American foulbrood, chalkbrood, and *Varroa destructor* mite infestation (Boecking and Spivak, 1999; Gilliam et al., 1988; Spivak and Reuter, 2001). Efforts to breed bees for hygienic behavior have been successful, and although a specific line of hygienic bees is available (Spivak and Gilliam, 1998), efforts are underway to help large-scale queen breeders select for hygienic behavior in their own stock. The more pests and diseases that can be controlled through bee behavior, the fewer chemical controls beekeepers need to use to treat pests and diseases. Not only is this beneficial in that it directly saves beekeepers money, it also lessens the probability of harmful interactions among treatments with other environmental factors, as noted above.

Resin collection by honey bees

Another social immunity trait that has great potential to help beekeeping is the colony's collection and deposition of plant resins in the nest. Bees collect resins from plants in their environment, and deposit them as 'propolis' (Simone-Finstrom and Spivak, 2010). Since the term 'propolis' is an apicultural term, in this dissertation I will refer to 'resin' as the material occurring on the plant or on bees' corbiculae (Fig 1) and refer to 'propolis' only as the material that has been deposited in the nest where it is typically mixed with wax. Much of the basic biology underlying this trait remains unexplored, although, in general, bees use resin as a building material, and feral honey bees will coat the entire inside surface of their nesting cavity with a layer of propolis (Fig 1) (Seeley and Morse, 1976). There is a great degree of natural variation among honey bee races in their tendency to accumulate propolis in the hive, from Caucasian bees which accumulate a relatively large amount to Italian bees that accumulate relatively little.

Figure 1 – Resin foraging by honey bees. Honey bee worker with resin on her corbiculae (left) and a cross-section of a feral honey bee nest in a tree cavity (right). The wood surrounding the comb is covered by a thin layer of dark propolis. Photos courtesy of Dr. Mike Simone-Finstrom and Gary Reuter.

The presence of propolis in the hive prevents chronic up-regulation of individual immune function in adult bees (Simone et al., 2009), which is practically important for beekeepers as this chronic up-regulation decreases colony productivity (Evans and Pettis, 2005). Honey bees that accumulate more propolis have also been reported to live longer and have healthier brood (Nicodemo et al., 2013). Unfortunately, bees deposit little resin in the conventional Langstroth hives used by most beekeepers compared to natural nesting cavities. The role of propolis against specific honey bee pathogens is largely unknown, yet honey bees slightly increase resin foraging upon challenge with the filamentous fungus *Ascophaera apis*, the causative agent of chalkbrood (Simone-Finstrom and Spivak, 2012). Resin use is prevalent among bees in the families *Apidae* and *Megachilidae*, including honey bees other than *A. mellifera*, stingless bees, and euglossine bees; however, the relationship between resin and health is obvious in a more distantly related Hymenopteran insect, the Swiss wood ant, *Formica paralugubris*. Their nest mounds contain 2.3 g of coniferous resin per liter of nesting material on average (Christe et al., 2002), and the presence of these resins in the nest increases the survival of both adults and larvae when challenged with entopathogenic bacteria and fungi

(Chapuisat et al., 2007). Based on these results, it appears plausible that propolis accumulation in honey bee nests can also increase honey bee resistance to specific diseases if resin is deposited in sufficient quantities. However, more research is needed to understand the specific mode of action of propolis, perhaps against specific diseases (e.g. contact vs. volatilization).

Honey bees utilize a variety of resinous plants, though we know little about the specific sources of bee-preferred resin as a whole. Plants secrete resin on young buds and leaves, stems, wounds, and fruits where it is collected by bees (Langenheim, 2003; Simone-Finstrom and Spivak, 2010). *Clusia spp.* and *Dalechampia spp.* are unusual in that they offer resin as a reward for pollination (Armbruster, 1984), while *Corymbia torelliana* is unusual in that it offers resin as a mechanism of seed dispersal (Wallace and Trueman, 1995). Most botanical sources of propolis have been discovered by direct observation and, to a limited extent, by chemical analysis (Bankova, 2006; Bankova et al., 2000; Crane, 1990). Major shifts in the botanical sources of resin can be viewed as a function of climatic region and the literature is mostly divided by temperate vs. tropical sources, though resin sources in cold and arid climates must also be considered.

Composition of resins utilized by bees

Resins are chemically complex and variable protective mixtures of phenolic and isoprenoid compounds (Langenheim, 2003), and this variability and complexity appears to make them a good defense against rapidly evolving pests and pathogens (Gershenzon and Dudareva, 2007; Langenheim, 2003). Unfortunately, the lack of clear acknowledgement that propolis is primarily a plant derived substance, and consequently a product of plant specialized metabolism that will vary based on the regional availability of specific resinous plant species, is a common omission in the propolis literature. This has led to mistaken views and generalizations concerning propolis biological activities (e.g. that propolis from a particular region is active against a particular pathogen, therefore all propolis can be expected to be active against this pathogen). Propolis activities need to be considered as a function of their botanical sources that are likely variable on a small regional scale, and perhaps even by the seasonality of resin collection

(Bankova et al., 1998) or developmental stage of the plant (Boege and Marquis, 2005; Valkama et al., 2004). There are no accounts of honey bees changing resin composition, except through the physical addition of other substances like wax.

It takes no more than a quick visual inspection to determine that propolis samples from different regions have different colors, textures, and smells that likely arise from their different plant sources and that they should be considered different plant-derived substances. The matter is further complicated by the fact that bees will forage from more than one resinous plant in an area, and in these cases the resulting propolis becomes a mixture of resins that may not be consistent over time. The unique mixture of compounds found in a propolis sample are, at best, restricted to propolis from a single apiary, though compound types indicative of major botanical sources (e.g. prenylated cinnamic acids) can sometimes be tracked over larger regions using GC-MS or HPLC (Bankova, 2006; Bankova et al., 2000). Still, it is very difficult to manually compare chromatograms and say definitively that a particular propolis sample originated from a particular plant, considering the typically large number of peaks generated and the presence of spurious peaks that come from minor botanical sources or contamination (see Fig 2 &3 in Park et al. 2003). The complexities that arise from sampling propolis in the hive could be avoided if sampling occurred on the level of individual resin foraging bees. In addition, if more unbiased analytical methods were utilized, like LC-based metabolomics, it would be easier to make a robust determination of resin botanical sources and, at the same time, determine characteristic compounds in these resins. Knowing the botanical sources of propolis on a large scale through the chemical tracking of individual resin-forager behavior allows for the distribution of bee-preferred compounds in the environment to be known. This would allow us to test how or why bees prefer some resinous plants over others, presumably through differing volatile components (Leonhardt et al., 2010), or how specific resinous plants in the environment impact bee health.

For the purpose of this dissertation, I will focus the discussion on plants occurring in the U.S., although resin foraging from a mentioned plant may have been observed elsewhere. It is thought that poplar resin is preferred by honey bees in temperate regions

when it is available, and bees are known to collect resin from *P. nigra* in Spain and central Europe (Tomás-Barberán and García-Viguera, 1993), and also *P. angustifolia, P. deltoides, P. fremontii, P. trichocarpa,* and *P. balsamifera* in the U.S, and *P. tremula* in Italy (Crane, 1990; Wollenweber and Buchmann, 1997). Trees in the *Populus* genus excrete large amounts of resin from secretory epidermal cells in leaf buds, but also from glands located at the teeth margins of leaves and at the base of the petiole depending on species (Fig 2) (Curtis and Lersten, 1974; Langenheim, 2003). Resins from different poplar species have been well characterized by GC-MS, and *Populus spp.* resins are mostly composed of flavonoid aglycones, chalcones, phenolic acids, and some sesquiterpenes (English et al., 1992, 1991; Greenaway and Whatley, 1991, 1990; Greenaway et al., 1991a, 1991b; Greenaway et al., 1988, 1987; Isidorov and Vinogorova, 2003). Interestingly, aspen resins (*Populus* section *Populus*) seem to mostly contain fatty acids and hydrocarbons (Greenaway et al., 1991c). A literature compilation of compounds discovered in U.S. poplar resins can be found in Appendix B.

Some species of birch (*Betula spp.*) can be important resin plants for honey bees in colder temperate climates, though not all birches secrete resin. Some birch species, such as European white birch (*Betula pendula*), silver birch (*B. platyphylla*), paper birch (*B. papyrifera*), and resin birch (*B. neoalaskana*), have glandular trichomes that collapse to form numerous resin droplets on juvenile stems (Palo, 1984; Rousi et al., 1991). Some species also secrete resin from foliar trichomes in buds (Palo, 1984; Valkama et al., 2003). Birch resins contain triterpenoids, like papyriferic acid, which function as important deterrents of mammalian herbivory, and also some flavonoids (Wollenweber and Dietz, 1981; Palo, 1984; Rousi et al., 1991; Taipale et al., 1993; Valkama et al., 2003). Alders, *Alnus spp.*, are closely related to birch and are also utilized by honey bees (Crane, 1990; Simone-Finstrom and Spivak, 2010). Like birch species, alders secrete resin from buds and young stems that contain primarily terpenoids with some phenolics and flavonoids (Wollenweber and Dietz, 1981; Langenheim, 2003).

Stone fruits, *Prunus spp.*, and buckeyes, *Aesculus spp.*, are abundant in wild and cultivated populations in the U.S., and bees have been observed foraging resin from apricot (*P. armeniaca),*sweet cherry (*P. avium*), and sour cherry (*P. cerasus*), and also

from horse chestnut (*A. hippocastanum*) (Crane, 1990). There is a lot of diversity in *Prunus* secretions, though *Prunus spp.* are generally known for the excretion of true gums from stem wounds, which contain only carbohydrates (Langenheim, 2003). The buds of some *Prunus spp.*, including sweet cherry and sour cherry, accumulate resins that contain flavonoid aglycones (Wollenweber and Dietz, 1981) and, more rarely, triterpenoids (Langenheim, 2003). Some *Aesculus spp.* also produce phenolic bud resins that contain flavonoid aglycones (Langenheim, 2003; Wollenweber and Dietz, 1981). Buckeyes are well recognized for the production of polyhydroxylated triterpenoid saponins, but it is not clear if these compounds also occur in their resins.

We know less about resin sources in arid climates compared to temperate and tropical regions, though several species are known to be targeted for resin collection by honey bees. Honey bees in Mexico forage resin from gum limbo, *Bursera simaruba* (Crane, 1990) which also occurs in Florida, while other members of the genus occur in Arizona and California. *Bursera spp.* are trees and shrubs that produce large amounts of terpenoid resins from trunk and stem wounds, and some species actually spray resin when injured (Becerra, 2001; Langenheim, 2003). *Bursera* resin is comprised mostly of mono- and sesquiterpenes, but also di- and triterpenes (Becerra, 2001). Chemical and observational studies of feral honey bees in the Sonoran Desert suggest that these bees collect resin from ragweed, *Ambrosia deltoidea*, and brittlebrush, *Encelia farinose* (Wollenweber and Buchmann, 1997). Ragweeds, *Ambrosia spp,*. are widespread in the U.S., and cause severe and prevalent seasonal allergies with their wind-blown pollens. In addition, ragweeds produce leaf-coating resins which cause contact dermatitis due to the presences of sesquiterpene lactones that have a methylene group on their lactone (Mitchell et al., 1971). These resins also contain some flavonoids with rare substitution patterns (Wollenweber et al., 1987). Brittlebrushes, *Encelia spp.*, are prevalent in the southwestern U.S. and have resinous stems and leaves. Benzopyrans and benzofurans are characteristic of brittlebrush resin (Proksch and Rodriguez, 1983), along with a diversity of methoxylated flavonoids (Proksch et al., 1988). Instead of benzopyrans and benzofurans, some *Encelia spp.* seem to produce sesquiterpene lactones in their resin (Srivastava et al., 1990). Honey bees also collect resin from creosote bush, *Larrea*

tridentate, in U.S. desert climates (Crane, 1990). Creosote bush is a dominant species in its native range and produces a leaf-coating resin with the distinctive smell of wood smoke from which the plant derives its common name (Langenheim, 2003). Its resin is mostly comprised of phenolic aglycones, mainly the lignin nordihydroguaiaretic acid, with some methoxylated flavonoid agylcones, mono-, and sesquiterpenes (Arteaga et al., 2005; Langenheim, 2003).

The flowers of *Clusia spp.* provide resin to bees in the tropical and subtropical Americas, and honey bees are known to forage from flowers of both *C. major* and *C. minor* (Tomás-Barberán and García-Viguera, 1993). As noted previously, *Clusia spp.* are one of the few plants that offer resin as a reward for pollination. The major components of *Clusia* resins are a distinctive class of photoactive phenolics known as benzophenones, and these compounds can be highly prenylated in *Clusia* (Tomás-Barberán and García-Viguera, 1993; Oliveira et al., 1996), though *Clusia* resin may also contain some triterpenes (Armbruster, 1984). Interestingly, *Clusia spp.* are mostly dioecious and distinct chemical differences have been shown between resins from male and female flowers of *C. grandiflora* (Lokvam and Braddock, 1999). Also, while resin from both male and female flowers were both antimicrobial, resins from female flowers were much better inhibitors of the bacterial bee pathogens *Paenibacillus larvae* and *P. alvei* (Lokvam and Braddock, 1999).

Honey bees have been observed foraging resin from lemon-scented gum eucalyptus, *Corymbia citriodora*, and candelabra tree, *Araucaria angustifolia*, in Brazil and mangoes, *Mangifera indica*, in India (Crane, 1990; Park et al., 2002); however, the distribution of these non-native genera are restricted to parts of California/Hawaii, Florida/Hawaii, and Florida, respectively, in the U.S. Eucalypts excrete large amounts of resin from wounds and fruits and are critical resin sources for stingless bees in Australia (Wallace and Trueman, 1995; Wallace and Lee, 2010). *Araucaria spp*. are unusual tropical/subtropical coniferous trees in that they excrete gum resins that contain carbohydrates in addition to specialized metabolites (Anderson and Munro, 1969). Analysis of *Araucaria* gum resins has revealed the presence of diverse lignans (Yamamoto et al., 2004; Aslam et al., 2013), but also flavonoids and diterpenes (Aslam et

al., 2013). Mangoes are commonly grown in gardens and for limited commercial production in Florida and also excrete a gum resin from wounds in both shoots and fruits. Mango resins contain xanthonoids, a distinct class of polyphenolic compounds, along with triterpenes, such as mangiferolic acid (Corsano and Mincione, 1965; Ghosal et al., 1978).

There are many other resinous plants supposedly utilized by bees, although the knowledge concerning both the biology and chemical composition of their resins is incomplete. It is not always clear if compounds reported from some plants come from internal tissues or resins, as this distinction is not the focus of many studies. The list of botanical sources of propolis in the U.S. provided by Crane, 1990, also includes ash (*Fraxinus excelsior*) and chestnut (*Castanea sativa*) which occur in the U.S. with many other members of their respective genera. Crane's list also reports resin collection from Brazilian peppertree (*Schinus terebinthifolius*), frangipani, (*Plumeria rubra*), and guava (*Psidium gaujava*) in Hawaii, and other members of these genera occur in California, Texas, Louisiana, and Florida. Honey bees are reported to collect resins from several other species in Brazil including five species of False willow (*Baccharis spp.*), bushmint (*Hyptis divaricata*), and rosewood (*Dalbergia ecastophyllum*) that do not occur in the U.S. (Daugsch et al., 2008; Park et al., 2002) Other *Baccharis spp.* than those mentioned in Park et al., 2002, and *Hyptis spp.* are common in the U.S., while other *Dalbergia spp.* can be found in Florida.

There are many other resinous plants that could possibly be utilized by bees for which there is not conclusive evidence, thus precluding a comprehensive discussion of each here (see Langenheim, 2003). While not discussed in depth in this dissertation, resin is very important to the biology of many bees other than *A. mellifera*, and many of the questions surrounding the health impact of resinous plants on native bees are very relevant areas of study. It is important to bear in mind that the native ranges of many of the reviewed resinous plant genera overlap, and we know relatively little about the preferences of honey bees between or among these genera. The general conclusion is that poplar resins are preferred when available, but there are many co-occurring species of poplar in the U.S., and it is unclear if or how honey bees make choices between

different resinous plants. It is also currently unclear if certain resinous plants are more beneficial to bees than others.

Utility of resin and propolis against specific honey bee disease

Resins are excreted by plants with the specific purpose of defending their tissues from predators, pathogens, and abiotic stressors (Langenheim, 2003), though the main benefit of resins to bees is thought to be derived from their antimicrobial properties. When propolis extracts are used to coat hive boxes in place of the natural propolis coating found in tree cavity nests, there is an overall decrease in bacteria present in the hive environment (Simone et al., 2009). It seems that propolis has a general antimicrobial activity, particularly against gram-positive bacteria (Kujumgiev et al., 1999; Banskota et al., 2001). Indeed, one of the most important honey bee diseases is caused by a gram-positive bacterium, *Paenibacillus larvae.*

P. larvae is the gram-positive, endospore-forming bacterium that causes American foulbrood (AFB) in honey bees. AFB can be very lethal to colonies, as ten or fewer endospores ingested by larvae 12-36 hrs after hatching is sufficient to cause mortality (Genersch, 2010). *P. larvae* is particularly troublesome because its endospores are viable almost indefinitely in typical environmental conditions and can contaminate beekeeping equipment and hive products for many years, facilitating the spread of AFB throughout apiaries (Genersch, 2010). Infected colonies (bees and equipment) are often burned to control the spread of this disease, resulting in significant economic loss for beekeepers (vanEngelsdorp et al., 2012). In the U.S, AFB can be suppressed through use of the antibiotics oxytetracycline and tylosin, though re-infection can occur through persisting endospores. Antibiotic treatment of AFB is also problematic due to the repeated development of oxytetracycline resistance by *P. larvae* in North America (Evans, 2003) and regulations that forbid honey production until six or four weeks post-treatment for oxytetracycline and tylosin, respectively (Reybroeck et al., 2012). No antibiotics are approved for treating AFB in the European Union (Reybroeck et al., 2012). Due to tightening regulations regarding the use and adaptation of antibiotics for

agricultural applications, it is unlikely that new antibiotic treatments for AFB will become available.

Propolis extracts inhibit the growth of *P. larvae in vitro* and may be a relatively non-toxic alternative to antibiotics, though propolis from different regions can be more or less effective (Bastos et al., 2008; Lindenfelser, 1967). Feeding bees or spraying comb with a propolis solution was able to halt AFB progression, though only a small number of colonies were used and re-infection occurred after treatments stopped (Lindenfelser, 1968). Treatments in that report were toxic to bee brood, though it was not clear if toxicity was caused by propolis or by solvent (Lindenfelser, 1968). Similar feeding and spraying experiments showed that propolis treatments reduce the number of *P. larvae* spores in honey from challenged colonies, but none of the control colonies in that report developed clinical symptoms of AFB for comparison (Antúnez et al., 2008).

The general antimicrobial activities of propolis may also be useful against *Ascosphaera apis*, the fungal ascomycete brood pathogen that causes chalkbrood in honey bees. As noted previously, honey bees respond to challenge by *A. apis* with slightly increased resin foraging (Simone-Finstrom and Spivak, 2012). Chalkbrood is found world-wide, but typically cannot destroy a colony like AFB (Aronstein and Murray, 2010); however, chalkbrood can significantly reduce colony productivity and be dangerous to weakened colonies (Aronstein and Murray, 2010). Like *P. larvae*, *A. apis* spores are persistent in the environment, can contaminate beekeeping equipment for long periods of time, and are only pathogenic if ingested by larvae (Aronstein and Murray, 2010; Hornitzky, 2001). Since there are no chemical agents efficacious in the control of chalkbrood, though a host of substances have been tested both in culture and in colonies, a management strategy that includes propolis could be a relatively non-toxic control mechanism that will not leave dangerous residues in the hive (Hornitzky, 2001). While there is some preliminary evidence that propolis can inhibit *A. apis* growth in culture (Babić et al., 2011), field research has shown that artificial enrichment of the hive with propolis decreases chalkbrood infection when colonies are challenged (Simone-Finstrom and Spivak, 2012).

The potential benefits of propolis against the natural enemies of bees are not restricted to its antimicrobial activity against bacteria and fungi. Surprisingly, propolis can be used either chemically in the lab or behaviorally by bees to kill or weaken several honey bee pests. Cape honey bees, *Apis mellifera capensis*, effectively control the nest parasitic beetle, *Aethina tumida*, where the beetles occur in their native African range by encapsulating these insects in propolis (Neumann et al., 2001; Neumann and Elzen, 2004). In contrast, *A. tumida* is an invasive pest in the southern U.S. that can quickly overrun healthy colonies, because Italian honey bees, *Apis mellifera lingustica*, encapsulate *A. tumida* at a much slower rate (Ellis et al., 2003; Neumann and Elzen, 2004). A targeted breeding effort for social encapsulation could lead to better control of this pest in the U.S. It has also been shown that the metabolism of the Varroa mite, *Varroa destructor*, is disrupted by direct application of propolis extracts *in vitro* (Garedew et al., 2002). Varroa infestation is arguably the single greatest threat to modern beekeeping, and colony-level research of the anti-Varroa activity of propolis could prove very useful. The naturally occurring monoterpene thymol is currently used as an impregnated strip-based fumigant for the control of Varroa, and many similar compounds occur in plant resins.

Though not shown specifically against any bee virus, propolis extracts also exhibit *in vitro* virucidal activity against some enveloped viruses (Amoros et al., 1992), while inhibiting the transmission and replication of the herpes and HIV viruses in cell culture (Gekker et al., 2005; Nolkemper et al., 2010). Bee virus transmission is facilitated within and between colonies by Varroa mites, and viruses may play an important role in large-scale colony losses in the U.S. (Cox-Foster et al., 2007). We are only beginning to understand the effects of viruses on honey bee health, though viruses may be a wide-spread and devastating disease issue. Natural and widely available methods for suppressing bee viruses, such as compounds derived from propolis, could be valuable to beekeepers.

While I have just provided substantial evidence supporting the potential specific benefits of resins and propolis to beekeeping, it is still unclear how these substances should be used by beekeepers to improve bee health in the field. Some honey bee

diseases are notoriously hard to scientifically control in the field (see Lindenfelser, 1968 and Antúnez et al., 2008), and careful research is needed to determine how propolis, resins, or any compounds derived from these substances should be used to benefit bees and beekeepers. It is important to note that some types of propolis may be dangerous to humans by contact (e.g. propolis made from *Schinus spp.*, *Mangifera spp.*, or *Ambrosia spp.* resin), so care must be taken when considering what 'kinds' of propolis or resins could be used for application. While bee-collected plant resins have positive effects on bee physiology (Simone-Finstrom et al., 2009; Nicodemo et al., 2013) and potential as inhibitors of pests and pathogens (see above), propolis should not be approached like a 'silver bullet' that will cure the complex problem of honey bee health. Compounds contained in various plant resins will be most effective as part of a series of management strategies that represent the best practices for a given environment. We should not ignore the potential benefits of social immunity traits related to propolis or the benefits derived from the presence of propolis in the hive, but there is still much research to be done to fully understand how propolis can be leveraged to benefit beekeepers.

Biologically active compounds isolated from propolis

Propolis has been used to treat a variety of human maladies since antiquity and is described in Egyptian hieroglyphs and many classical texts (Kuropatnicki et al., 2013). Ancient Egyptians were known to use propolis in the embalming process, while the Greeks, and later the Romans, utilized the antimicrobial and healing properties of propolis to treat wounds (Kuropatnicki et al., 2013). Tinctures of propolis are still used widely today as topical antiseptics, healing agents, dental treatments, and sore throat remedies in continental Europe, Brazil, Japan, and Russia. For example, propolis extracts are commonly used in modern Brazilian veterinary medicine to prevent or treat infections in open wounds, surgical wounds, and sores in place of antibiotics with good clinical results (F. Wilson, personal communication). It has been reported that propolis from different regions has antimicrobial, hepatoprotective, anti-inflammitory, immune system modulation, and anti-proliferation activities; however, propolis is not utilized in modern Western medicine for any of these purposes (Marcucci, 1995; Burdock, 1998; Sforcin,

2007; Sforcin and Bankova, 2011). There seem to be barriers to human clinical studies of propolis activities against disease, probably because propolis is a mixture of many differentially active compounds that can be difficult to standardize (Bankova, 2005a; Sforcin and Bankova, 2011).

Though the clinical use of propolis as a mixture is lacking, many compounds with anti-proliferation and antimicrobial properties have been isolated from propolis samples. Caffeic acid phenyl-ethyl ester (CAPE), a phenolic acid ester commonly found in poplar-type propolis, is a potential antitumor agent and chemopreventative reported to interfere with the oxidative activation of tumor cells, suppress tumor growth, and promote tumor cell death in many cancers *in vitro* (Banskota et al., 2001b; Burdock, 1998). PMS-1, a clerodane-type diterpene, is another potential anticancer agent isolated from Brazilian propolis that was shown to inhibit or kill human hepatocellular carcinoma HuH13 cells and human lung carcinoma cells *in vitro* (Banskota et al., 2001b; Burdock, 1998). Several other propolis-derived compounds with *in vitro* anticancer activities have been isolated from Brazilian propolis including two other clerodane-type diterpenoids, (Z)-2,2-dimethyl-8-prenyl-benzopyran-6-propenoic acid (a singly prenylated diterpene), methyl caffeic acid (a methylated phenolic acid), dihydrokaempferide (a flavonoid), nemorosone and its methylated derivatives (polyisoprenylated benzophenones), and artepillin C (a diprenylated p-coumaric acid) (Burdock, 1998; Banskota et al., 2001; Bankova, 2005b). To the author's knowledge, only two propolis-derived compounds have shown any anti-proliferation activity *in vivo*. Artepillin C was shown to cause tumor cell death when injected into human carcinoma and malignant melanoma tumors transplanted into mice, while oral doses of CAPE decreased the number of lung tumor nodules in rats (Banskota et al., 2001b; Sforcin and Bankova, 2011).

Though, to the author's knowledge, there have not been any *in vivo* tests of propolis-derived compounds against infectious disease, many studies have reported *in vitro* antimicrobial activities against bacteria and fungi. The antimicrobial properties of propolis are generally attributed to phenolics, polyphenolics, prenylated phenolic acids, and diterpenes (Bankova, 2005b). Many common phenolic and polyphenolic compounds occur in propolis, with some being generally recognized for their antimicrobial activity.

Simple phenolic acids are common in poplar-type propolis with cinnamic acid, ferulic acid, p-coumaric acid, gallic acid, and caffeic acid having known activities against bacteria, fungi, and/or viruses (Marcucci, 1995; Cowan, 1999). The common flavonoids apigenin, chrysin, galangin, kaempherol, quercetin, naringenin, and pinocembrin have also all been found in propolis and all have documented antimicrobial activities (Marcucci, 1995; Cushnie and Lamb, 2005). Specifically, chrysin was found to bind bacterial extra-cellular proteins, while quercetin and naringenin were found to inhibit bacterial motility (Mirzoeva et al., 1997; Cowan, 1999). In addition, quercetin and apigenin were reported to inhibit bacterial DNA gyrase, while galangin was reported to disrupt bacterial membranes (T. P. T. Cushnie and Lamb, 2005).

Propolis-derived prenylated phenolic acids, diterpenes, and polyisoprenylated benzophenones also have reported activities, though these compounds are not as common as phenolic acids and flavonoids in propolis. Aga et al. and Marcucci et al. isolated several polyprenylated p-coumaric acids from Brazilian propolis that likely originated from *Baccharis spp.* and showed strong activity against three different bacteria and *Trypanosoma cruzi*, the parasitic euglenoid that causes Chagas' disease (Aga et al., 1994; Marcucci et al., 2001). Bankova et al. reported four antimicrobial labdane-type diterpenic acids isolated from Brazilian propolis, including isocupressic acid, acetylisocupressic acid, imbricatoloic acid, and communic acid, that likely originated from *Araucaria spp.* (Bankova et al., 1996). More recently, Popova et al. reported nine antimicrobial terpenes (including seven diterpenes and two cycloartane triterpenes) isolated from Greek *Cupressaceae* or *Pinaceae* propolis active against a diverse panel of 11 bacteria and fungi (Popova et al., 2009). Two polyisoprenylated benzophenones, nemorosone and guttiferone A that are abundant in *Clusia spp.* propolis showed strong activity against the malaria parasite, *Plasmodium falciparum* (Monzote et al., 2011)

Banskota et al. displayed the diversity of active compounds when they isolated three prenylated p-coumaric acids, two methoxy flavonoids, two diterpenes, and two diterpenic acids from the sample propolis sample that all showed high activity against three strains of the gastric ulcer-causing bacterium, *Helicobacter pylori* (Banskota et al., 2001a). This diversity has proven to be both a curse, when considering the use of crude

propolis in clinical medicine, and a blessing, when considering the potential to discover new biologically active compounds in propolis. Biological activity data is most powerful when coupled to knowledge of the active compound's botanical source, as mentioned previously. Doing so increases the value of propolis as a hive product by allowing beekeepers to make rational landscape decisions about apiary environment that help to maximize and standardize useful resin-derived compounds in the propolis produced by their bees.

Summary of research objectives

The research presented in this dissertation has three primary objectives:

1) Develop MS-based metabolomic methods for tracking resin foraging behavior in honey bees, *Apis mellifera*.
2) Determine the diversity in antimicrobial activity of U.S. propolis against the bee pathogens *P. larvae* and *A. apis*.
3) Isolate compounds responsible for activity against *P. larvae* and *A. apis* in active propolis samples from the U. S.

As noted previously, determining the botanical sources of propolis is critical to understanding the biology of resin use by honey bees and promoting its utility and application. The limiting factor in this determination is often the difficultly involved in either observing resin foraging behavior or using traditional chemical analyses on propolis sampled from colonies. In Chapter 2 I present and utilize highly-sensitive and discriminatory MS-based metabolomics methods that can determine the botanical origins of resins collected by individual resin-foraging bees, while also characterizing the chemical and biological differences between different *Populus spp.* resins occurring in the U.S.

The diversity in antimicrobial activity of U.S. propolis against bee pathogens has been largely unexplored. Given the many different U.S. climatic regions with lots of comparative botanical diversity, and the fact that we do not known how different botanical sources of resin impact bee health, determining if any regions harbor botanical

conditions for making exceptionally antimicrobial propolis is important. In Chapter 3 I present an evaluation of regionally diverse propolis samples, covering the desert southwest to the Atlantic coast, for antimicrobial activity against the bacterial and fungal bee pathogens *P. larvae* and *A. apis*.

Understanding the specific compounds responsible for propolis antimicrobial activity is key to the standardization and widespread application of propolis in humans and animals. As the general physiological function of resin in plants is to provide a stable and chemically complex barrier against infection, propolis has proven to be a source of novel and biologically active compounds. Isolation of biologically active compounds from propolis is also an easy, sustainable way to exploit plant compounds that would otherwise require intense harvesting to obtain. In Chapter 4, I present the isolation and characterization of novel and known compounds against *P. larvae* and *A. apis* from active propolis samples identified in Chapter 3

Chapter 2

Metabolomics Reveal the Origins of Antimicrobial Plant Resins Collected by Honey Bees[1]

Summary

The deposition of antimicrobial plant resins in honey bee, *Apis mellifera*, nests has important physiological benefits. Resin foraging is difficult to approach experimentally because resin composition is highly variable among and between plant families, the environmental and plant-genotypic effects on resins are unknown, and resin foragers are relatively rare and often forage in unobservable tree canopies. Subsequently, little is known about the botanical origins of resins in many regions or the benefits of specific resins to bees. We used metabolomic methods as a type of environmental forensics to track individual resin forager behavior through comparisons of global resin metabolite patterns. The resin from the corbiculae of a single bee was sufficient to identify that resin's botanical source without prior knowledge of resin composition. Bees from our apiary discriminately foraged for resin from eastern cottonwood (*Populus deltoides*) and balsam poplar (*P. balsamifera*) among many available, even closely related, resinous plants. Cottonwood and balsam poplar resin composition did not show significant seasonal or regional changes in composition. Metabolomic analysis of resin from 6 North American *Populus spp.* and 5 hybrids revealed chromatographic peaks characteristic to taxonomic nodes within *Populus*, while antimicrobial analysis revealed that resin from different species varied in inhibition of the bee bacterial pathogen, *Paenibacillus larvae.* We conclude that honey bees make discrete choices among many resinous plant species, even among closely related species. Bees also maintained fidelity to a single source during a foraging trip. Furthermore, the differential inhibition of *P. larvae* by resins from *Populus spp.*, which are thought to be preferential for resin

[1] Published as Wilson MB, Spivak M, Hegeman AD, Rendahl A, Cohen JD (2013) Metabolomics reveals the origins of antimicrobial plant resins collected by honey bees. PLoS One 8(10): e77512. doi:10.1371/journal.pone.0077512

collection in temperate regions, suggests that resins from closely related plant species may have different benefits to bees.

Introduction

Honey bees, *Apis mellifera*, are highly social insects that live in large colonies (e.g., 50,000 individuals). One cost of social living is an increased rate of disease transmission among individuals, and honey bees are highly prone to a diverse set of pathogens and parasites (Schmid-Hempel, 1998). Managed populations of honey bees in the U.S. are in serious decline, and there has been a 61% decrease of registered colonies from 1947 to 2008 (vanEngelsdorp and Meixner, 2010). This decrease is due, in large part, to unsustainable winter losses (vanEngelsdorp and Meixner, 2010; vanEngelsdorp et al., 2012) caused by the combined effects of diseases, parasites, pesticides, and nutritional deficiencies (Cox-Foster et al., 2007; vanEngelsdorp et al., 2009; Alaux et al., 2010; Johnson et al., 2010; van Dooremalen et al. 2012). This is particularly alarming because honey bees are estimated to contribute $15-20 billion dollars annually to U.S. agriculture from pollination services, their major contribution to the U.S. agricultural economy (Johnson, 2010). Beekeeping practices and regulatory issues indicate that the most sustainable solutions to problems plaguing bees will be derived from promoting their natural defenses through breeding and habitat enhancement. While honey bees have only 1/3 of the genes involved in individual immunity compared to the solitary insects *Drosophila* (fruit fly) and *Anopheles* (mosquito) (Evans et al., 2006), they do have a suite of cooperative behaviors that contribute to colony health called 'social immunity' (Cremer et al., 2007; Evans and Spivak, 2009; Wilson-Rich et al., 2009). An example of social immunity is hygienic behavior, where honey bees work together to detect and remove diseased brood from the nest, resulting in colony-level resistance to pathogens and parasites (Gilliam et al., 1988; Spivak and Reuter, 2001; Evans and Spivak, 2009; Wilson-Rich et al., 2009)

In addition to hygienic behavior, bees also deposit antimicrobial plant resins in their nests that have important immunological benefits (Simone et al., 2009) (Fig. 1). Feral honey bees coat the entire inside surface of their nesting cavity with resin (Seely

and Morse, 1976), but managed honey bees deposit comparatively little resin in conventional beekeeping hive boxes. Resins are complex mixtures of phenolic and isoprenoid compounds (Langenheim, 2003) secreted by plants to provide protection against predators and pathogenic microorganisms (Arrhenius and Langenheim, 1983; Langenheim and Hall, 1983). The chemical composition of resins is complex and variable within and among plant families, traits that makes resin production a good defense against rapidly evolving pests and pathogens (Witham, 1983). Many organisms, including insects, birds, and humans, collect and use resins to protect against their own pathogens and parasites (Ghisalberti, 1976; Chapuist et al., 2007; Mennerat et al., 2009).

Honey bees collect resins on their hind legs and deposit them in the nest where the resin, often mixed with wax, is called 'propolis' by apiculturists. Honey bees are known to collect resin from a wide variety of plants depending on what is available in their environment, though bees in temperate climates are thought to collect mainly from *Populus* (cottonwood, poplar, aspen), but also from *Betula* (birch), *Salix* (willow), *Alnus* (alder), and *Aesculus* (horse chestnut) (Alfonsus, 1933; Ghisalberti, 1976; Crane, 1990; Bankova et al., 2000). The botanical origin of resin is important because propolis harvested from colonies in different climatic regions, and thus from different botanical sources, could vary in its antimicrobial properties. Indeed, it was shown that propolis samples from different regions do vary in their ability to inhibit *in vitro* growth of the bee pathogen, *Paenibacillus larvae* (Bastos et al., 2008). This effect is most likely due to the diversity in specialized metabolites secreted by the resinous plants available to bees in different regions; however, it seems that propolis has a general inhibitory effect on gram-positive bacteria and fungi (Kujumgiev et al., 1999). This should be expected as the general inhibition of microorganisms is a role resins play in plant defense (Levin, 1976; Arrhenius and Langenheim, 1983).

Identifying the botanical sources of resins collected by honey bees can be challenging since resin foraging is relatively infrequent compared to pollen foraging (Simone-Finstrom and Spivak, 2012), the variation in resin among and between plants is mostly unknown, and foraging can occur in the canopy of trees. The botanical sources of propolis remain a mystery in most regions of the world, though 35 plant families with 88

genera that contain known resinous species occur in the continental U.S. (Langenheim, 2003; USDA-NRCS, 2013). Traditional chemical analysis has been somewhat successful in identifying botanical sources of bee-collected resin by sampling at the colony level (Bankova et al., 2006). However, these methods are difficult to apply to generally unknown, variable, and complex substances, like resin, due to the amount of *a priori* information required. The exact identity of a characteristic signature compound must be known, chemically analyzed, and available as a standard. One or more of these requirements are often missing. Traditional analysis is also inefficient at describing biological variation among large numbers of samples, which is key to uncovering subtle differences among complex mixtures. Colony-level sampling is also problematic because bees collect resin from more than one plant and mix them in the hive.

In order to identify the botanical sources of bee-collected resins and measure their species specific and seasonal variation without any prior knowledge of resin composition, we used metabolomic methods as a type of 'environmental forensics' to track resin forager behavior on the level of individual bees. We also used antimicrobial assays to explore potential differences in the derived benefits of collecting resin from some plants relative to resin from other plant species.

Materials and Methods

Sampling

Honey bees (*Apis mellifera lingustica*) were sampled from a single apiary located on the University of Minnesota, St. Paul campus. Resin was dissolved directly off the corbiculae of individual bees with acetonitrile. Resin extracts from each bee were diluted to 10% acetonitrile in water for analysis but were not normalized due to limited amounts of material.

Resin was sampled from wounds and buds of individual plants, dissolved in acetonitrile, and diluted to 1.5 mg/mL in 10% acetonitrile for analysis. Multiple wounds were sampled, if available, while six buds per individual plant were sampled.

Plant resins were collected from plants growing on the St. Paul MN campus and from plants established in the greenhouse from hardwood cuttings. Greenhouse was

maintained at 25°C without supplemental light. Plants were watered biweekly and fertilized once with Osmocote Plus (Scotts Company – Marysville, OH).

Data collection

Spectral data for Fig. 2-5 were generated using high-performance liquid chromatography (HPLC) (UltiMate 3000, Thermo-Fisher Scientific, Waltham, MA) coupled to Fourier transform mass spectrometry (FTMS) (Q-Exactive, Thermo) operated at 17,500 resolution in full scan, (-) ionization mode. Gradient: water-acetonitrile, column: Zorbax XDB C_{18} (Agilent Technologies, Santa Clara, CA), 2.1 x 100 mm, 1.8 µm particle size, flow rate: 350 µL/min. Metabolic fingerprints in Fig. 3, 4, and 5 were generated using ultra-performance liquid chromatography (UPLC) (Acquity LC, Waters, Milford, MA) coupled to time-of-flight (TOF) MS (Waters LCT Premier XE) in both (+) and (-) ionization modes. Gradient: water-acetonitrile, column: Waters BEH C_{18}, 1.0 x 100 mm, 1.8 µm particle size, flow rate: 130 µL/min. Metabolic fingerprints in Fig. 6. were generated using UPLC (Waters Acuity) coupled to TOF-MS (G2 Synapt, Waters) in both (+) and (-) ion mode. Gradient: water-acetonitrile, column: Zorbax Eclipse XDB C_{18} (Agilent), 2.1 x 100 mm, 1.8 µm particle size, flow rate: 350 µL/min.

Data analyses

All data analyses were performed using a custom R script developed in our laboratory (Appendix C). Automated peak discovery in raw MS data was performed with the R package XCMS (Smith et al., 2006; Tautenhahn et al., 2008) using the *centWave* algorithm. Parameters used: ppm = 10, peakwidth = c(5,50), fitgauss = TRUE, verbose.columns = TRUE. Peaks were grouped across samples using XCMS's *nearest* algorithm. Parameters used: mzCheck = 2, rtCheck = 5. Positive and negative ion mode peaks were combined into a single matrix for statistical analysis. A quality control sample containing equal amounts of each biological sample was run every ten samples, and spectral peaks that did not appear in all technical replicates of this quality control sample were eliminated from the analysis.

Bacterial inhibition assay

P. larvae (NRRL # B2605, USDA Agricultural Research Service culture collection) was grown in brain/heart infusion broth (Difco) fortified with 1 mg/L thiamine HCl. Growth inhibition was measured spectrophotometrically (Specta Max 190, Molecular Devices), using a 96 well-plate growth assay, as the percent OD_{600nm} of treated wells relative to untreated control wells after six hours of incubation and shaking at 37°C.

Results

Resin metabolite diversity available to bees

To discern what options bees have for foraging targets, the resinous plant diversity within common foraging range (3.2 km) of our St. Paul, Minnesota, campus apiaries was determined. The dominant resin-producing species in the area was *Populus deltoides* (eastern cottonwood), but *Pinus spp.* (pine) and *Picea spp.* (spruce) were also common. In addition, there were small stands of *Populus balsamifera* (balsam poplar) and hybrid poplars of unknown parentage, and scattered *Abies balsamea* (balsam fir), *Pseudotsuga menziesii* (Douglas fir), *Larix laricina* (tamarack larch), and *Aesculus hippocastanum* (horse chestnut) in the study area. *Populus tremuloides* (American aspen), *Populus grandentata* (big tooth aspen), and *Salix sp.* (willow) occurred at least once, and may be resinous in other locations (Crane, 1990), but were not obviously resinous at the time of sampling. It is important to note that all species were mostly interspersed among each other in the study area, though groups of *P. deltoides* and *P. balsamifera*, and individual *Pinus strobes, Pinus syvestris,* and *A. balsamea* were closest in proximity to the apiary. Sampled resins were compositionally complex with both qualitative similarities and differences, especially among genra (Fig. 2-5).

It is not known why bees forage for specific resins in the field. To explore the possibility of antimicrobial activity as a criterion for resin preference, the *in vitro* activity of local resins against *Paenibacillus larvae*, a bee brood pathogen, was measured. Resin from different species varied in their ability to inhibit *P. larvae* (Table 1), with resin from *P. gluaca* being the most inhibitory, achieving complete growth inhibition at 0.05

mg/mL. *A. hippocastanum* and *P. sylvestris* resin did not completely inhibit *P. larvae* growth within the experimental concentration range, though their resins did show some inhibition of growth (data not shown).

Using metabolomic forensics to reveal the botanical sources of resin

Twenty six individual resin foraging bees, typically carrying ≤ 5 mg of resin, were captured returning to two colonies over three sampling events in July 2011. Captured bees collected dark red and bright yellow resins, which match the visual description of resin from *P. balsamifera* (red), and *P. deltoides* (yellow), or hybrid poplars of unknown variety (yellow) occurring in the area. *P. deltoides*, *P. balsamifera*, and hybrid poplars occurring within two miles of the experimental apiary were sampled in June and July 2011 by washing resinous buds with acetonitrile. Resin was also collected from *P. deltoides* near Jamestown, ND, in July 2011 to test for regional variation. It is unclear how the environment impacts the expression of resin metabolites in *Populus spp.*, though it has been reported that increased light intensity does increase leaf resin accumulation in tropical *Hymenaea* and *Copaifera* species without effecting resin composition (Langenheim et al., 1981). Jamestown, ND is ~510 km northwest of the St. Paul study site and has a significantly different landscape (urban vs. prairie/wetland) with a slightly drier and cooler climate.

Metabolite "fingerprints" of resin samples collected from bees and plants were acquired by reversed-phase HPLC time-of-flight (TOF) mass spectrometry in both positive and negative ion modes. A quality control sample was created by pooling equal amounts of all biological samples (thus theoretically containing all peaks that could occur in the resulting dataset) and injected after every 10 samples throughout the analytical run to act as a reference for automated peak detection. Spectral peaks were discovered in the analytical data with the R package XCMS (Smith et al., 2006; Tautenhahn et al., 2008) using the *centWave* algorithm and matched across different samples using the *nearest* algorithm. Only peaks discovered in all technical replicates of the quality control sample were included in our analysis (313 spectral peaks), to ensure that only high confidence peaks would be used in the statistical analysis. This data analysis strategy was directed

toward the discovery of unique and exclusive metabolites among samples and not focused on differences in concentrations or proportions. Principal components analysis (PCA) was used to summarize metabolite patterns in the data (Fig. 6) (Trygg et al. 2007). Sensitivity testing showed that the grouping pattern of the PCA scores plot was not greatly affected by performing different scaling transformations ($\log_{(X)}$) of the peak intensity data, which influences the importance of peak intensity differences relative to the presence/absence of peaks among samples in the analysis.

Bee-collected resin and plant-collected resins group together in PCA space, with 10 bees foraging from *P. deltoides*, 15 bees foraging from *P. balsamifera*, and no bees foraging from any of the hybrid poplar populations (Fig. 6)[1]. The regional variation between *P. deltoides* resin collected in Jamestown, ND vs. resin collected in St. Paul, MN was negligible.

Seasonal variation in P. deltoides and P. balsamifera resin

Propolis composition has been reported to change seasonally in some regions (Bankova et al., 1998). This could be due to changes in bee foraging behavior, where bees change their preference for some resin plants over others over the course of the season. Alternatively, seasonal changes in propolis composition could be due to seasonal changes in resin availability or composition, such as changes in plant resin flow or changes in plant specialized metabolism.

We sampled *P. deltoides* and *P. balsamifera* around the experimental apiary throughout the 2011 resin collection season (May, June, July, August, and October) to test if detectable changes in resin composition occurred. Data acquisition and analysis were performed as described in the previous section. Fig. 7 shows a gradual shift in *P. balsamifera* resin composition by month. The data agree with our visual observations that *P. balsamifera* resin changes from yellow/orange in the active growing season to dark red when buds begin to set for the winter. Out of 382 *P. balsamifera* spectral peaks,

[1]One bee foraged from an unknown source that did not match the patterns of any previously sampled plant species, and was eliminated from the analysis (data not shown).

only one verified peak was exclusively found in all October samples (*m/z* = 483.409, RT = 23.53 min., positive ion mode). 13 other verified peaks were identified as possible late season indicators (Table 2).

The seasonal changes observed in *P. deltoides* resin were simpler than those observed for *P. balsamifera* resin. Two distinct groups formed, resin collected in October vs. resin collected in all other months (Fig. 8). Out of 352 *P. deltoides* spectral peaks, two verified peaks were found exclusively in all October samples (*m/z* = 339.220, RT = 13.15 min.; *m/z* = 295.220, 12.62 min., both negative ion mode). Seven other verified peaks were identified as possible late season indicators (Table 2).

Difference in resin metabolites and biological activity among Populus spp.

The genus *Populus* is widely regarded as a preferential source of resin for honey bees in temperate climates (Crane, 1990; Bankova et al., 2000, Langenheim, 2003). Clearly, this is true at the apiary examined in this study (Fig. 6), although it is not known why *Populus* is preferable to other resin producers or if there are preferences among *Populus spp*. Studies focused on phenolic compounds extractable with diethyl ether and analyzed as their trimethylsilyl derivatives have shown that resins from North American *Populus spp.* can have different compositions (Greenaway et al., 1987; English et al., 1991; English et al., 1992).

To test the diversity of metabolite composition and bee-relevant antimicrobial activities of North American poplar resins, 11 species and hybrids were propagated via hardwood cuttings in the greenhouse for analyses (*P. balsamifera, P. angustifolia, P. trichocarpa, P. nigra, P. deltoides, P. fremontii, deltoides* x *trichocarpa, trichocarpa* x *deltoides, deltoides* x *nigra, deltoides* x *maximowiczii*, (*deltoides* x *trichocarpa*) x *trichocarpa*). Resin was harvested and analyzed as described above. Only one *P. deltoides* cutting survived, so the analysis was supplemented with 5 samples of *P. deltoides* resin collected in the study area used previously. Fig. 9 summarizes the compositional relatedness found among the different species/hybrids. Of 344 spectral peaks, several exclusive peaks were characteristic of *P. trichocarpa, P. angustifolia,* or *P. deltoides x maximowiczii* respectively (Table 3), and each species/hybrid had a unique

combination of non-exclusive spectral peaks that appeared in all resin samples from a given species/hybrid. Most species/hybrids did not have any exclusively characteristic spectral peaks; however, many peaks were found to be characteristic of terminal phylogenic nodes (Hamzeh and Dayanandan, 2004) within *Populus* (*P. deltoides*/P. *fremontii*, *P. angustifolia*, *P. balsamifera*/*P. trichocarpa* - Table 3).

As the range of many abundant North American *Populus spp.* overlap, bees are commonly presented with a choice between closely related resin producing plants. To test the diversity in potential benefit of North American poplar resins to bees, resin extracted for analytical analysis was also tested for antimicrobial activity against *P. larvae*. Fig. 10 clearly shows differences in antimicrobial activity among the different species with *P. trichocarpa* being the most and *P. angustifolia* the least inhibitory.

Discussion

Honey bees in the study region have many resin sources to choose from that are diverse in metabolite content and antimicrobial activity (Fig. 2-5), and it appears they favor some plants over others (Fig. 6). Strong resin plant preference has been shown in stingless bees (Tribe: Meliponini) (Leonhardt and Bluthgen, 2009; Leonhardt et al., 2010) as well as a slight male over female *B. dracunculifolia* preference (~10%) by honey bees in Brazil (Teixeira et al., 2005).

To our knowledge, this is the first example of using chemical analysis to track individual resin forager behavior. Importantly, these methods were sensitive enough to track resin foraging behavior by analyzing the resin from a single leg of an individual bee. Metabolomic forensics confirmed that honey bees collected resin from *P. balsamifera* and *P. deltoides* (Fig. 6), but not from the numerous hybrid poplars located within common foraging range. We confirmed that honey bees have a high fidelity to a single botanical source of resin during a single foraging trip, even when multiple closely related species occur in very close proximity that are also active foraging targets (Fig. 6).

Regional environment had little effect on *P. deltoides* resin composition when comparing resin sampled near Jamestown, ND, and in St. Paul, MN (Fig. 6), while season had only a subtle affect on overall resin composition, mostly among resins collected in

October compared to all earlier months (Fig. 7 & 8). The relevance of these subtle compositional changes in relation to bee activity is uncertain, as all foraging stops by mid-September in the study region due to the seasonal decrease in temperature. At least in this study area, large changes in propolis composition by season would probably be due to changes in resin availability or forager preference, and not by gross shifts in specialized plant metabolism.

Resin sampled from greenhouse grown North American poplars showed that some metabolites were unique to taxonomic nodes, with resin from closely related species being surprisingly different (Table 3, Fig. 9). Hybridization did not seem to produce novel metabolites. *P. deltoides x trichocarpa* and *P. trichocarpa x deltoides* resins had different intermediate compositions compared to their parental species (Fig. 9), which may indicate some maternal effects on resin metabolism. Due to the clear appearance of taxonomic node-specific metabolites, future work could provide insight into the evolution of specialized metabolism in the *Populus* genus. Overall, this metabolomics approach provided a powerful method to discriminate among patterns of resin metabolites from closely related species and hybrids, while also identifying specific metabolites that were characteristic signatures among species/hybrid groups.

Collected resins varied in antimicrobial activity against the highly infectious brood pathogen, *P. larvae* (Table 1, Fig. 10). Although the study colonies were not infected with this pathogen, the data suggest that availability, proximity, and perhaps toxicity may play roles in the selection of resins by bees. *P. deltoides* and *P. balsamifera* were targets of resin foraging and also the closest abundant species to our experimental apiary. North American poplars differentially inhibited the growth of *P. larvae* and Fig. 10 shows that even among plant species in the same genus, a bee's choice of resin could have profound consequences for their ability to reduce the overall microbe load within the nest cavity and prevent or fight off disease. *P. balsamifera* resin was more strongly inhibitory than *P. deltoides* resin, but there was no obvious preference for the more inhibitory resin among captured resin foragers (Fig. 6 & Fig. 10). Future studies with deeper sampling might uncover aspects of preferences not obvious with the sample size used in this study.

As many North American poplars commonly co-occur in the same environment (*P. deltoides* and *P. balsamifera* in Minnesota, for example), bees often have to choose between closely related resin sources and it is not well understood how bees locate preferential resin sources. Resin foraging frequency does increase in several bee species under certain conditions. Stingless bees increase resin foraging in response to ant attacks, while honey bees increase resin foraging when intentionally exposed to the larval fungal pathogen *Ascophaera apis*, the cause of chalkbrood (Simone-Finstrom and Spivak, 2012). It would be informative to determine if bees change resin preference, along with frequency, during these events. Resin foraging may also be a tactile response to crevices and rough textures within the hive (Leonhardt and Bluthgen, 2009; Simone-Finstrom et al., 2010), and roughing the inside surface of standard bee boxes might encourage the deposition of more resin.

Chemical analysis is a highly informative alternative to observational studies of resin foraging, as direct observation of resin foraging is difficult and prone to error. Metabolomic methods uniquely allow for the analysis of many samples without targeting specific signature metabolites. We were able to compared samples using > 300 LC-MS chromatographic peaks in over 100 samples and summarize their differences during a single analysis (Fig. 9), which could not be accomplished with traditional analytical methods. Many of the peak differences detected were not obvious and would likely have been missed with only visual inspection of the raw LC-MS data.

We seek to understand the botanical sources and biological activities of resins in the field and how resin foraging behavior changes in response to environmental factors, such as infection and other biological stresses. If we can discover plants with preferable and more antimicrobial resins in different regions, it should be possible to create better environments that promote bee health by supporting behaviors and managerial strategies that lead to natural disease resistance.

Acknowledgements

We would like to thank A. David, B. McMahon, T. O'Brien, M. Zinkgraf, G. Howe, S. Kolpak, S. DiFazio, and M. Howard for providing hardwood cuttings of

Populus spp.; S. Harvey and J. Dalluge for technical LC-MS help; R. Olivarez for providing packaged bees; G. Reuter for managing the bees; G. Gardner and F. Gleason for reading, commenting on, and discussing the manuscript; members of the Boswell/Cohen/Gardner/Hegeman lab group for discussions and input.

Figure 1 – Honey bee resin collection. Top-down view of standard beekeeping equipment displaying a resin forager with red resin attached to her hind legs. Managed honey bees deposit resin mainly at the hive entrance, inner cover, and on top of the movable frames.

Figure 2 – Resin metabolite diversity of studied angiosperms. Base-peak negative-ion chromatogram ("fingerprints") of resin collected from individual *Populus spp.* (poplar) and *Aesculus hippocastanum* (horse chestnut) trees within 2 miles of the study apiary.

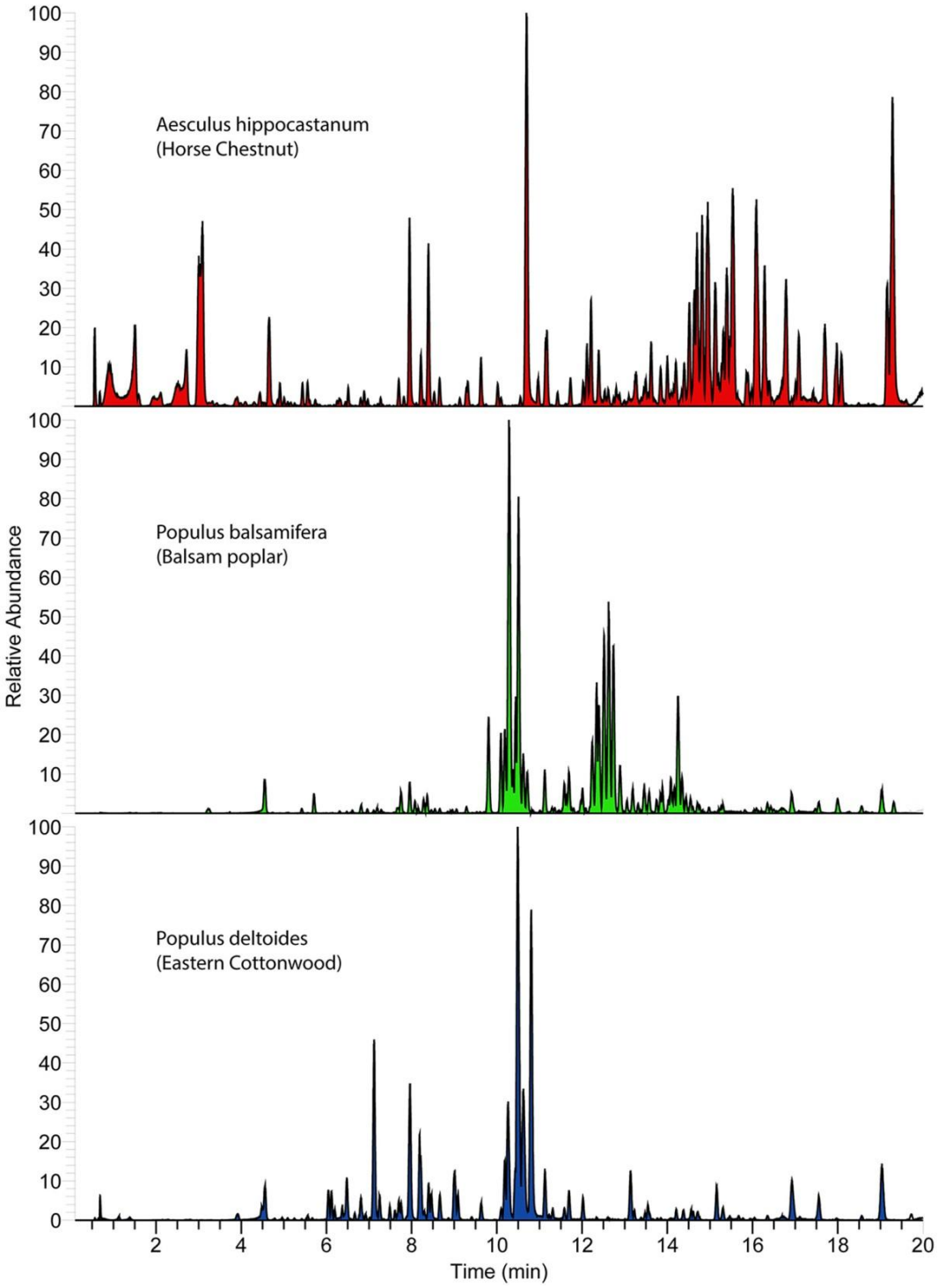


Figure 3 – Resin metabolite diversity of studied *Pinus sp.* (Pine). Base-peak negative-ion chromatogram ("fingerprints") of resin collected from individual *Pinus spp.* within 2 miles of the study apiary.

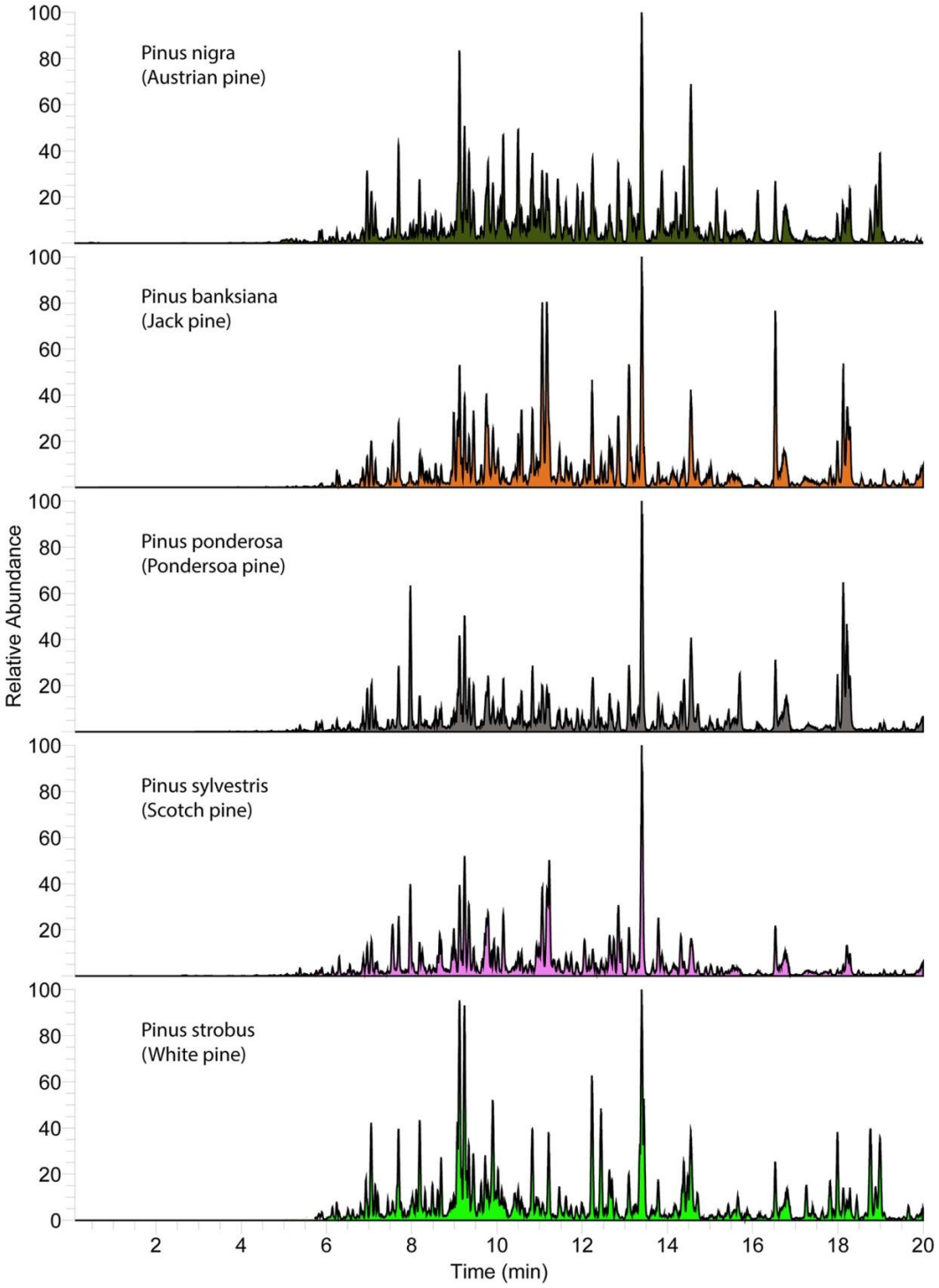


Figure 4 – Resin metabolite diversity of studied *Picea sp.* (Spruce). Base-peak negative-ion chromatogram ("fingerprints") of resin collected from individual *Picea spp.* within 2 miles of the study apiary.

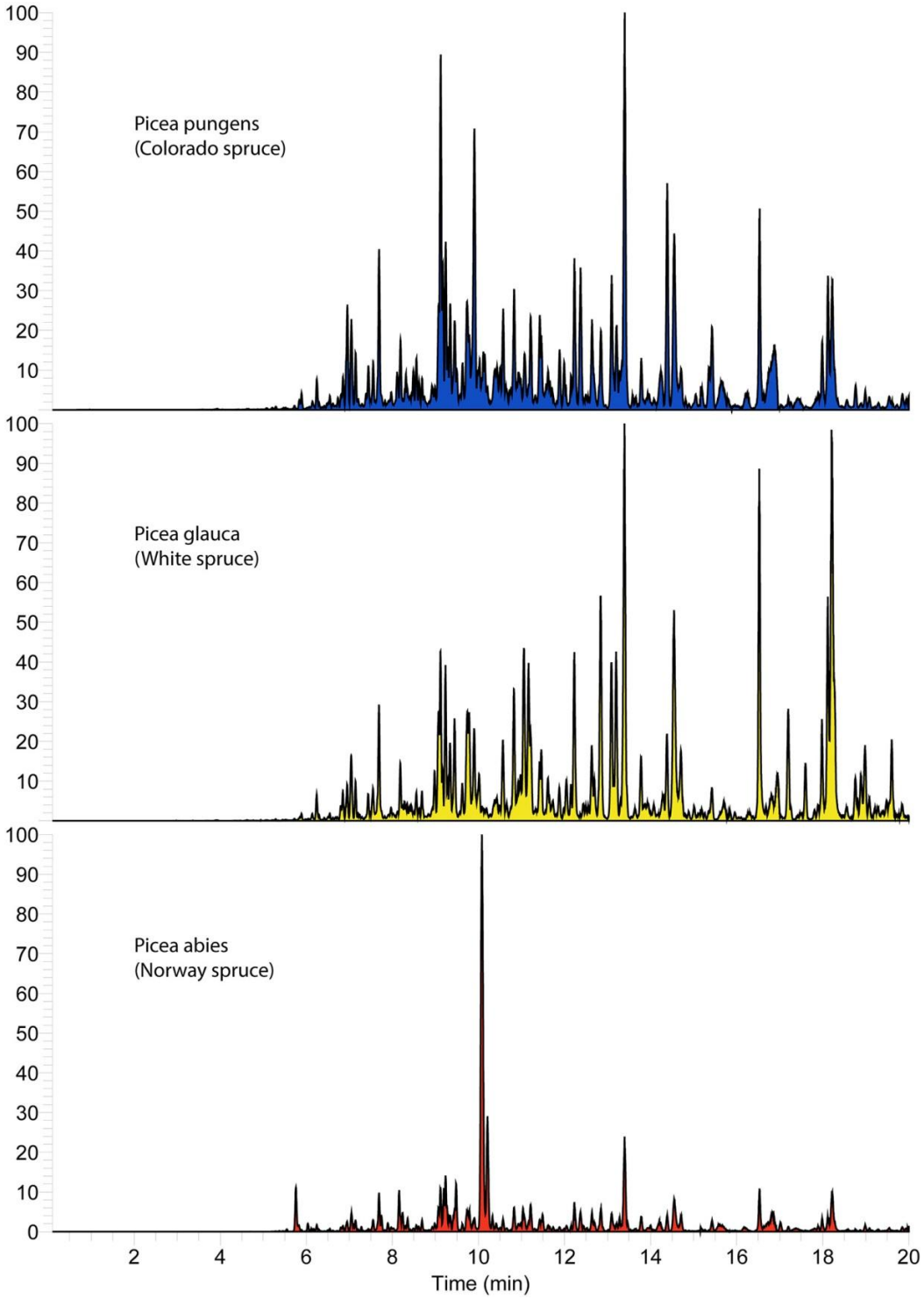


Figure 5 – Resin metabolite diversity of other studied conifers. Base-peak negative-ion chromatogram ("fingerprints") of resin collected from other conifers within 2 miles of the study apiary.

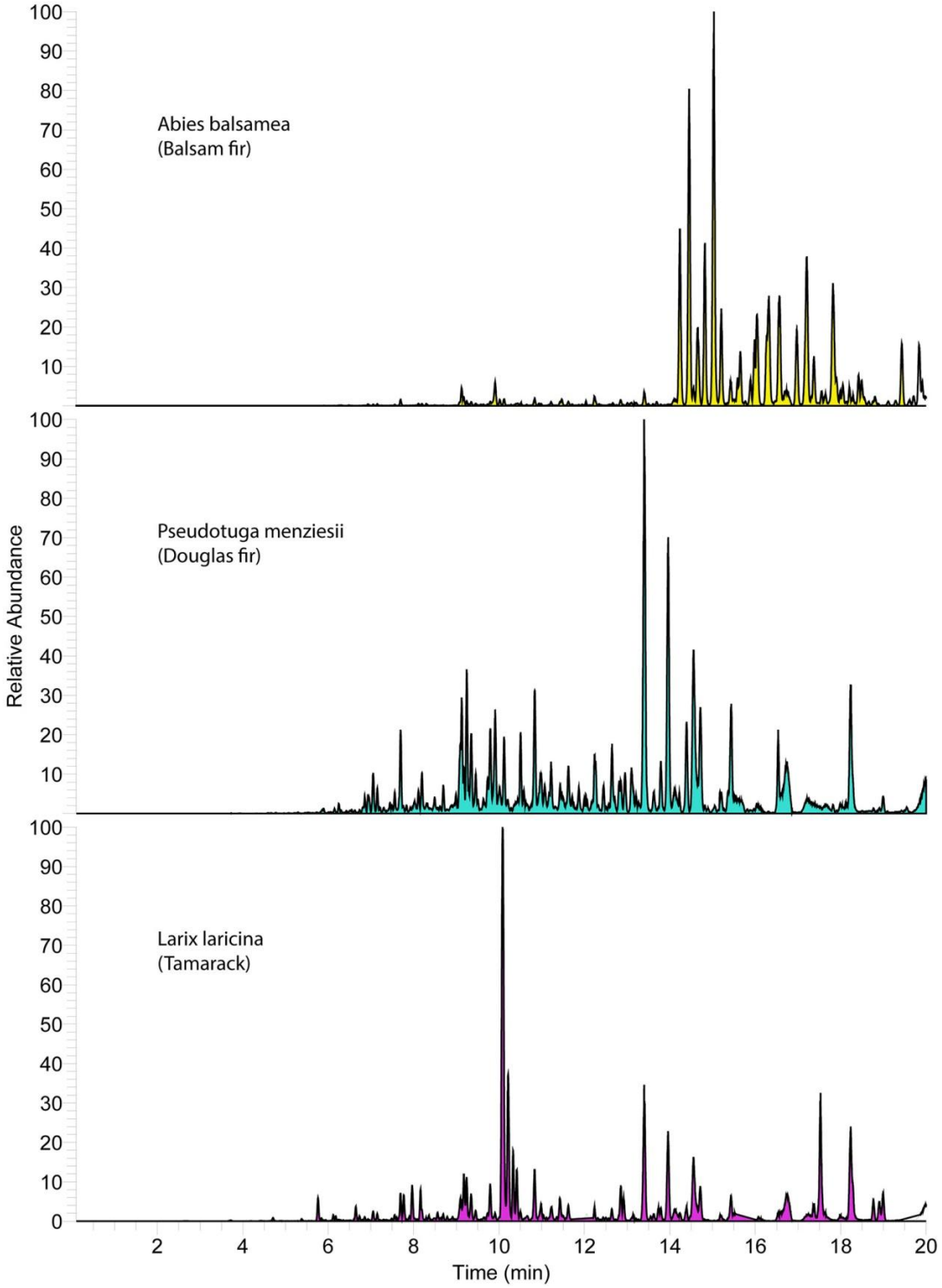


Figure 6 – Honey bees collect resin from *P. balsamifera* and *P. deltoides*. (Top) Example of a *P. deltoides* resin metabolite 'fingerprint'. (Bottom) PCA scores plot of resin 'fingerprints' from individual resin-producing plants and honey bee resin foragers. Each point represents the spectral composition of a biological sample. Points that are closer together have more spectral peaks in common than with points that are farther apart. 54.35% of the total variation in the data set is shown. Hybrid poplars were sampled once in June, while *P. deltoides* (Eastern cottonwood) and *P. balsamifera* (balsam poplar) once in June and once in July. N = 25 for resin from foraging bees, N = 11 for *P. balsamifera*, N = 16 for *P. deltoides* (9 from near Jamestown, ND), and N = 5 for each poplar hybrid population.

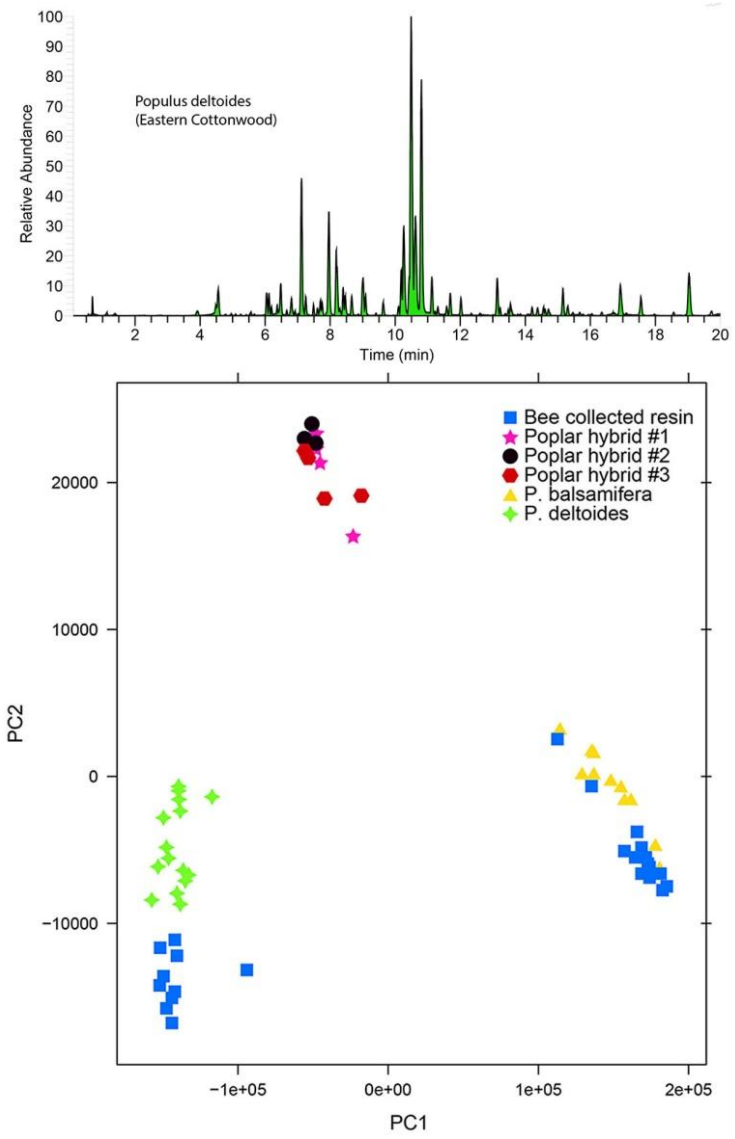


Figure 7 – Seasonal variation in *P. balsamifera* resin. PCA scores plot of resin 'fingerprints' from individual *P. balsamifera* plants collected throughout the growing season. 43.99% of the total variation in the data set is shown. N = 4 individuals in May, N = 5 individuals in June, August, and October.

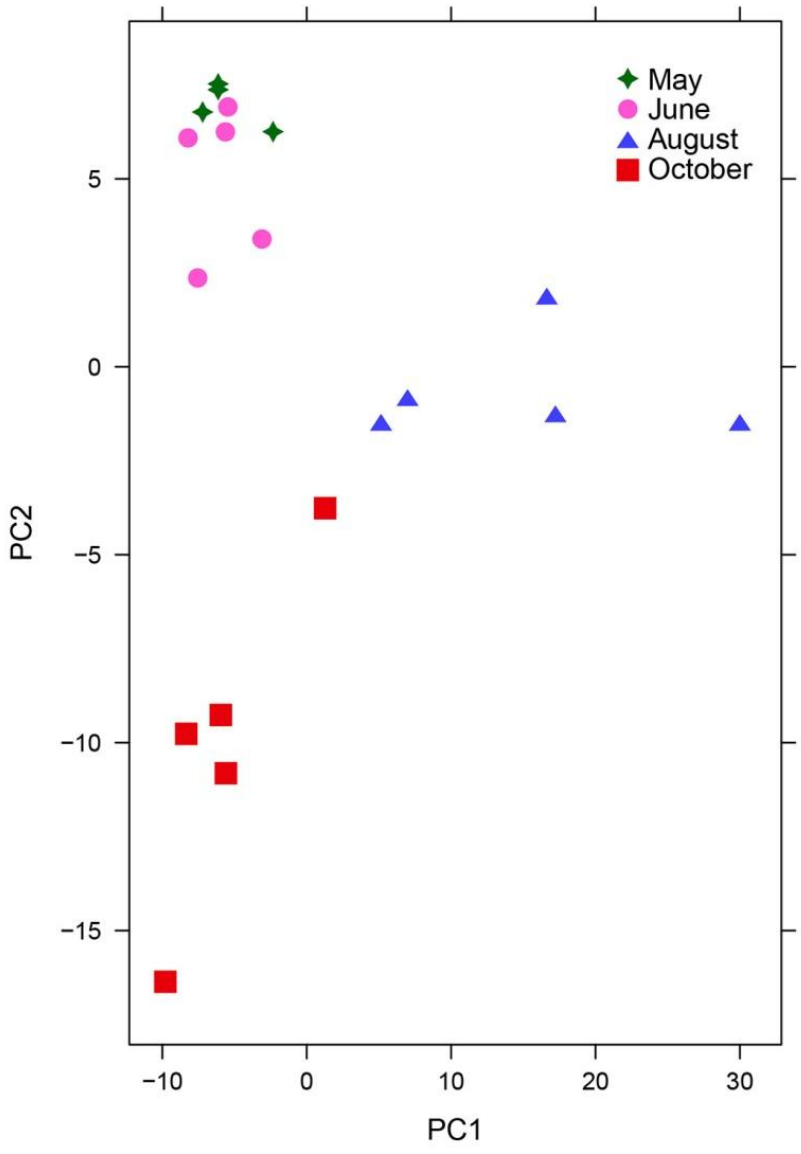


Figure 8 – Seasonal variation in *P. deltoides* resin. PCA scores plot of resin 'fingerprints' from *P. deltoides* plants collected throughout the growing season. 40.86% of the total variation in the data set is shown. N = 2 individuals in May, N = 5 individuals in June, N = 9 individuals in July (from near Jamestown, ND), N = 5 individuals in August, N = 5 individuals in October.

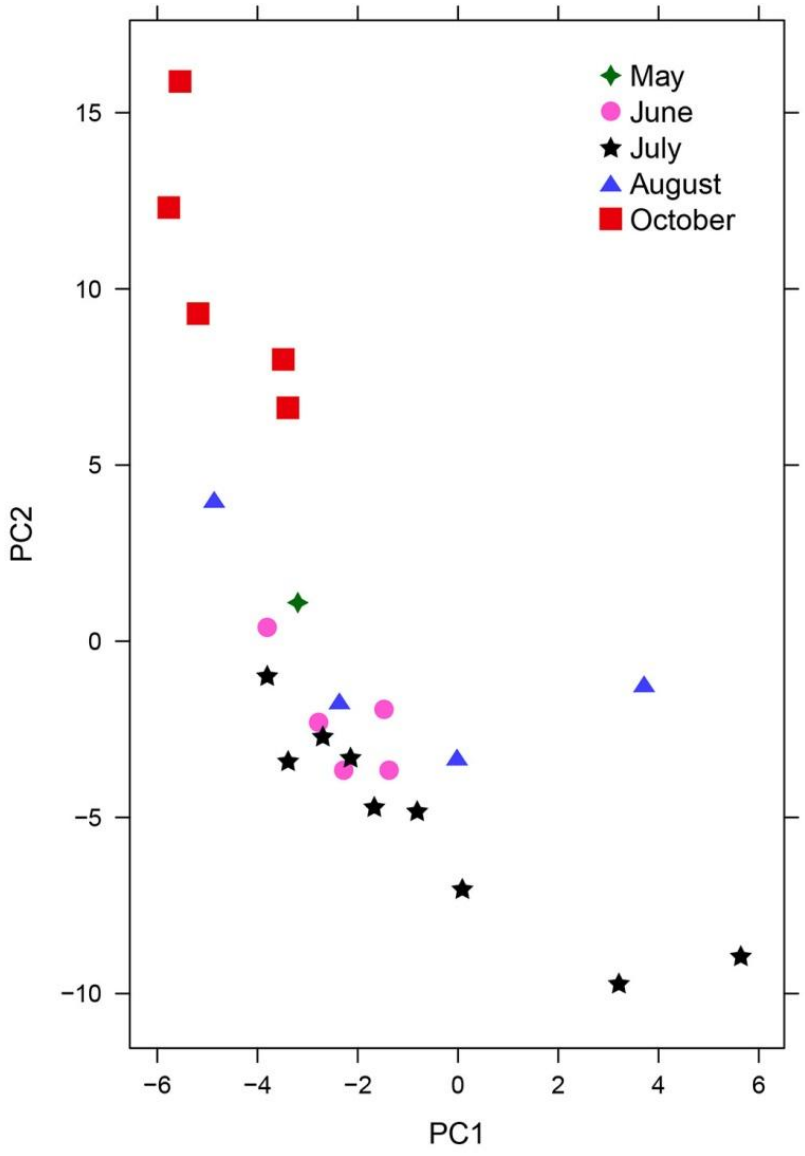


Figure 9 – Compositional differences in *Populus spp.* resin. PCA scores plot of resin 'fingerprints' from 11 different *Populus spp.* and hybrids grown under greenhouse conditions. Pure species are indicated by closed shapes, while hybrids are indicated by open shapes. 49.11% of the total variation in the data set is shown. d x m = *P. deltoides x maximowiczii* (N = 8), d x n = *P. deltoides x nigra* (N = 4), d x t = *P. deltoides x trichocapra* (N = 6), t x d = *P. trichocarpa x deltoides* (N = 5), (t x d) x d = *P. (trichocarpa x deltoides) x deltoides* (N = 14). N = 6 for *P. deltoides* and *P. nigra*, N = 12 for *P. fremontii*, N = 14 for *P. trichocarpa*, N = 18 for *P. angustifolia*, N = 5 for *P. balsamifera*.

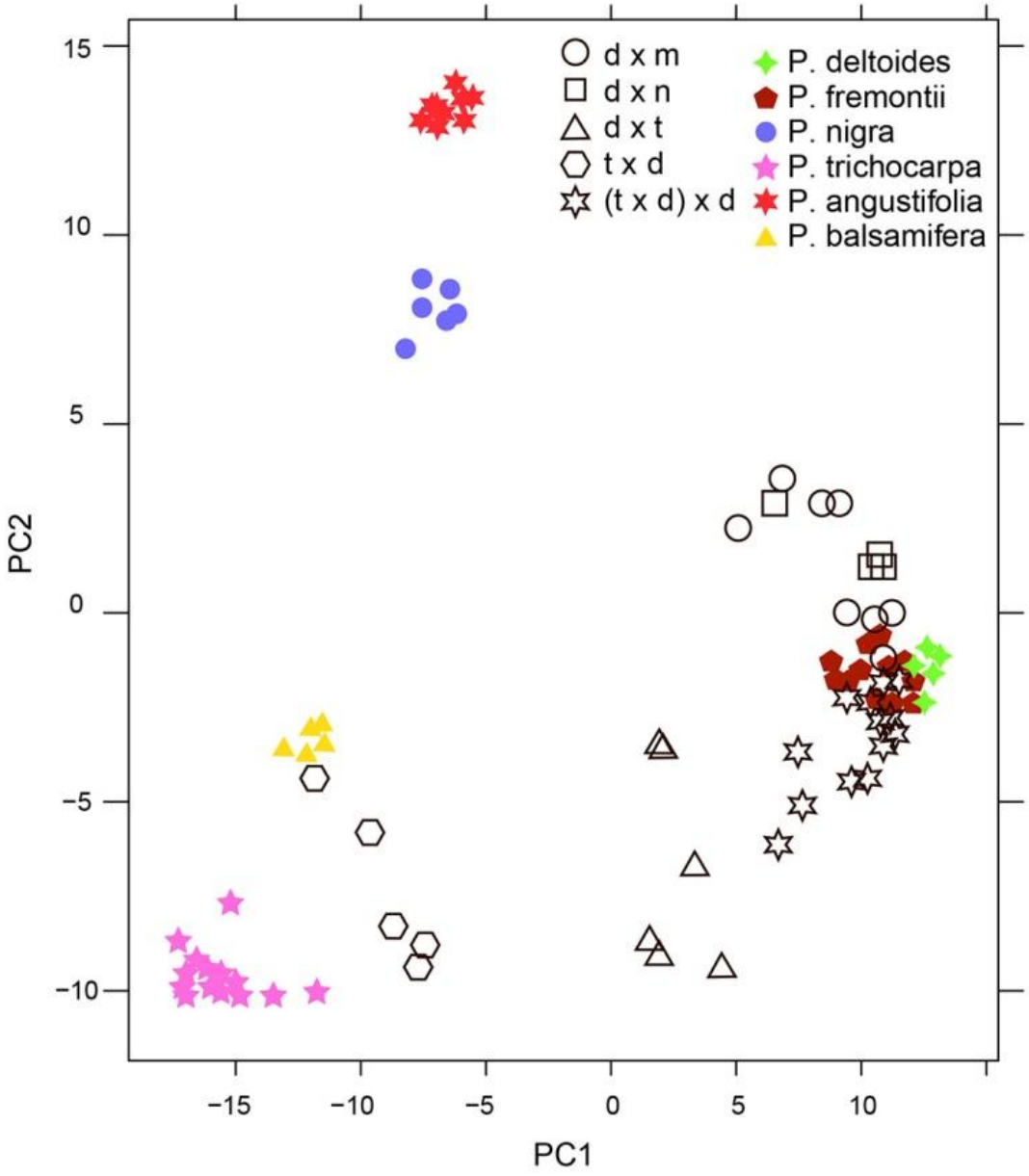


Figure 10 – Inhibition of *Paenibacillus larvae* by *Populus spp.* resin. Semi-log plot of the antimicrobial activity of resin from six representative *Populus spp*. individuals against *P. larvae*, a brood pathogen of honey bees. Antimicrobial activity was evaluated spectrophotmetrically at OD_{600} relative to untreated controls. N = 8 replicates per sample per concentration of resin.

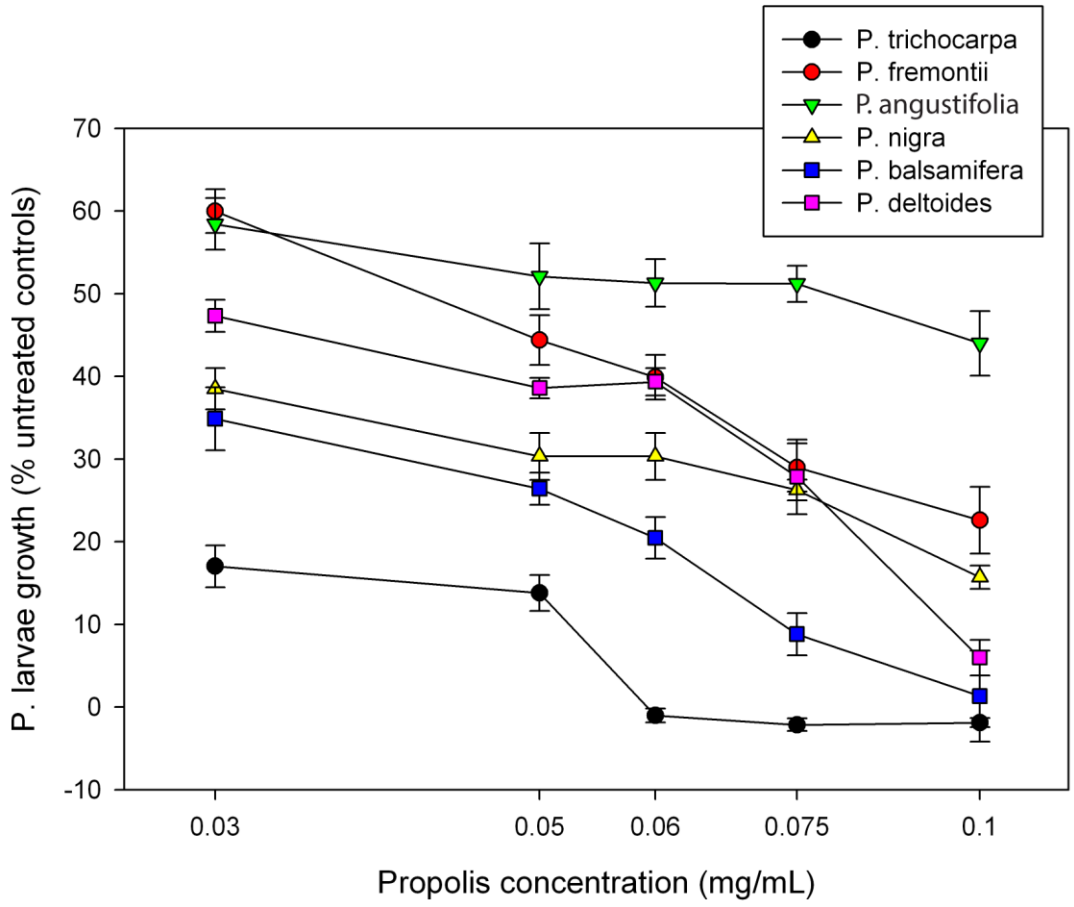


Table 1 – Inhibition of pathogen growth by local resins. Table describes the concentration at which the bee pathogen, *Paenibacillus larvae*, was completely inhibited by resin collected from local plants in a spectrophotometric growth assay (≤ 1% OD_{600} of untreated controls).

Resin Source	Minimum inhibitory concentration
Picea glauca	0.05 mg/mL
Larix laricina	0.06 mg/mL
Pinus banksiana	0.06 mg/mL
Pinus ponderosa	0.06 mg/mL
Populus balsamifera	0.075 mg/mL
Picea abies	0.1 mg/mL
Pinus nigra	0.1 mg/mL
Abies balsamea	0.125 mg/mL
Picea pungens	0.125 mg/ml
Pinus strobus	0.125 mg/mL
Pseudotsuga menziesii	0.175 mg/mL
Populus deltoides	0.175 mg/mL
Aesculus hippocastanum	> 0.175 mg/mL
Pinus sylvestris	> 0.175 mg/mL

Table 2 – Spectral peaks representing late-season indicators in *P. balsamifera* and *P. deltoides*. Table arranged by peak appearance in a given sample group. The number of samples within a month group in which each peak appears is indicated in the last column. Retention time was rounded to the nearest 0.01 min. Mass accuracy was 5-10 ppm. Rt = retention time.

Mass (m/z)	Retention time (min)	Ion Mode	Appearance (by month)
P. balsamifera			
483.409	23.53	+	5/5 Oct
353.099	11.07	-	4/5 Oct
621.271	14.57	-	4/5 Oct
595.205	14.85	-	4/5 Oct
491.288	14.57	+	1/5 Aug, 5/5 Oct
475.247	14.57	-	1/5 Aug, 5/5 Oct
531.279	14.57	+	1/5 Aug, 4/5 Oct
461.272	14.83	+	1/5 Aug, 4/5 Oct
625.212	14.57	-	1/5 Aug, 4/5 Oct
653.208	14.57	-	1/5 Aug, 4/5 Oct
519.276	15.77	-	1/5 Aug, 4/5 Oct
537.273	16.27	-	1/5 Aug, 4/5 Oct
521.292	15.98	-	3/5 Aug, 5/5 Oct
373.291	19.6	-	1/5 June, 5/5 Oct
P. deltoides			
339.220	13.15	-	5/5 Oct
295.220	12.52	-	5/5 Oct
293.220	13.15	-	4/5 Oct
295.239	13.13	+	4/5 Oct
503.401	18.35	+	1/4 Aug, 5/5 Oct
519.356	16.25	-	1/4 Aug, 5/5 Oct
525.384	18.35	+	2/4 Aug, 5/5 Oct

Table 3 – Spectral markers of terminal taxonomic nodes in *Populus*. Table arranged by peak appearance in a given species group. The number of samples within a species group in which each peak appears is indicated in the last column. Retention time was rounded to the nearest 0.01 min. Pd = *P. deltoides*, Pf = *P. fremontii*, Pa = *P. angustifolia*, Pb = *P. balsamifera*, Pt = *P. trichocarpa*, dm = *P. deltoides x maximowiczii*. Mass accuracy was 2 ppm. Rt = retention time.

Mass (m/z)	Retention time (min)	Ion Mode	Appearance (by node or species)
Taxonomic Section *Aigeiros*			
291.0634	6.33	+	6/6 Pd, 12/12 Pf
563.1664	6.83	+	6/6 Pd, 12/12 Pf
395.1100	6.95	+	6/6 Pd, 12/12 Pf
307.0584	7.77	+	6/6 Pd, 12/12 Pf
241.0866	7.77	+	6/6 Pd, 12/12 Pf
269.0825	7.77	+	6/6 Pd, 12/12 Pf
329.1025	7.77	+	6/6 Pd, 12/12 Pf
357.1330	10.10	+	6/6 Pd, 12/12 Pf
379.1147	10.10	+	6/6 Pd, 12/12 Pf
273.0770	12.10	+	6/6 Pd, 12/12 Pf
267.0644	6.33	-	6/6 Pd, 12/12 Pf
327.0863	7.80	-	6/6 Pd, 12/12 Pf
313.0737	9.60	-	6/6 Pd, 12/12 Pf
Taxonomic Section *Tacamahaca*			
363.1214	9.40	+	18/18 Pa
339.1215	9.40	-	18/18 Pa
121.0654	9.10	+	5/5 Pb, 14/14 Pt
301.1074	10.85	+	5/5 Pb, 14/14 Pt
303.1239	11.30	+	5/5 Pb, 14/14 Pt
421.1654	11.40	+	5/5 Pb, 14/14 Pt
121.0652	11.30	+	5/5 Pb, 14/14 Pt
301.1083	11.40	+	5/5 Pb, 14/14 Pt
553.2219	12.40	+	5/5 Pb, 14/14 Pt
301.1086	13.00	+	5/5 Pb, 14/14 Pt
285.1126	13.40	+	5/5 Pb, 14/14 Pt
553.2221	13.40	+	5/5 Pb, 14/14 Pt
315.1225	15.20	+	5/5 Pb, 14/14 Pt
405.1182	4.30	-	5/5 Pb, 14/14 Pt
287.0912	9.10	-	5/5 Pb, 14/14 Pt
389.1389	11.60	-	5/5 Pb, 14/14 Pt
551.2079	12.40	-	5/5 Pb, 14/14 Pt
521.1958	12.50	-	5/5 Pb, 14/14 Pt
433.1644	13.10	-	5/5 Pb, 14/14 Pt
404.1544	13.40	-	5/5 Pb, 14/14 Pt
535.2126	15.30	-	5/5 Pb, 14/14 Pt
437.2103	6.40	+	14/14 Pt
551.1533	6.70	+	14/14 Pt
161.1306	9.00	+	14/14 Pt
135.0412	2.80	-	14/14 Pt
Taxonomic Section *Aigeiros/Tacamahaca*			
209.0802	9.10	-	8/8 dm

Chapter 3

Regional variation in composition and antimicrobial activity of U.S. propolis

Summary

Propolis is a substance derived from antimicrobial plant resins that honey bees use in the construction of their nests. Propolis use in the hive is an important component of honey bee social immunity and confers a number of positive physiological benefits to bees. There is evidence that resins are also part of bees' natural defenses against at least one pathogen, but it is unknown how the diversity of antimicrobial activities among resins might impact bee health. The work described in Chapter 2 indicates that resins from closely related *Populus spp.* have differing levels of activity against a bacterial bee pathogen, *Paenibacillus larvae*, and thus certain resins may be more or less beneficial to bees. In this research, the variation in antimicrobial activity of propolis from 12 climatically diverse regions in the U.S. against two bee pathogens, *P. larvae* and *Ascophaera apis*, is characterized. Samples differed greatly in their ability to inhibit bacterial and fungal growth, and propolis from Fallon, NV, Beaumont, TX, and Aspen, CO displayed high activity against both pathogens. Antimicrobial assays and analytical analysis revealed that the observed differences in antimicrobial activity were not due to antimicrobial compounds previously isolated from propolis. Metabolomic analysis of regional propolis samples revealed that each sample was compositionally distinct, as spectra from each sample contained a unique number of shared and exclusive peaks. Surprisingly, spectra from the most active propolis samples shared a large number of peaks, even though they originated from very different botanical landscapes. Propolis from Aspen, CO, Tuscon, AZ, and Raleigh, NC, contained large numbers of exclusive peaks, which might indicate that these samples originated from relatively unique botanical sources.

Introduction

Social immunity traits are cooperative behaviors among social insects that contribute to disease resistance on the colony level (Cremer, 2007; Wilson-Rich et al. 2009; Evans and Spivak, 2010). Resin collection has been demonstrated as a form of social immunity in ants and honey bees and is related to a number of positive physiological benefits in these species (Chapuist et al., 2007; Simone et al., 2009). These benefits manifest in honey bees, *Apis mellifera*, as increased adult longevity, increased brood vitality, and the prevention of chronic immune system up-regulation leading to decreased productivity (Evans and Pettis, 2005; Simone et al., 2009; Nicedemo et al., 2013). These benefits make resin collection an important aspect of bee biology for beekeepers in general.

Evidence also suggests that resins play a role in the natural defenses of ants and bees against specific pathogens. Nest enrichment with resin improved the survival of wood ant larvae challenged with the ant bacterial pathogen *Pseudomonas fluorescens* (Chapusiat et al., 2007); however, the effects of resins or propolis against bee bacterial pathogens, such as *Paenibacillus larvae*, in whole colonies remain unclear (Lindenfelser, 1968; Antúnez et al., 2008). *In vitro* studies have shown that propolis from several regions can inhibit *P. larvae* growth (Lindenfelser, 1967; Bastos, et al., 2008). To this end, several compounds have been isolated from Bulgarian propolis with reported *in vitro* activity against *P. larvae* including benzyl ferulate, pentenyl ferulate, pinocembrin, pinobanksin-3-acetate, and a mixture of caffeate esters (Bilikova et al., 2012). In addition to role of resins against bacterial pathogens, honey bees respond to challenge with the bee pathogenic fungi *Ascophaera apis* by increasing resin foraging (Simone-Finstrom et al., 2013), while enrichment of colonies with propolis or resins actually prevents *A. apis* and *Metarhizium anisopliae* fungal infections in honey bees and wood ants, respectively (Chapusiat et al., 2007; Simone-Finstrom et al., 2013). It is unknown how active compounds in resins or propolis interact with pathogens in the colony.

All races of honey bees collect plant resins from their environment and deposit them in their nesting cavity as propolis, though some races deposit relatively little compared to others (Seely and Morse, 1976; Crane, 1990; Simone-Finstrom and Spivak,

2010). For clarity, we refer to 'resin' as the material occurring on the plant or on a bee's corbiculae, and refer to 'propolis' as the material that has been deposited in the nest and typically mixed with wax or other substances. Propolis is used as a nest building material in natural and manmade bee hives, and feral honey bees use it to coat the entire inside surface of their nesting cavities (Seely and Morse, 1976). In contrast, many honey bees deposit comparably little propolis inside smooth, manmade bee boxes, but instead use propolis to glue down frames and covers. This behavior has likely led to selection against bee stock that deposit large amounts of propolis, as this can make managing the hive difficult for beekeepers (Fearnly, 2001). We now know that this selection for easy management may have come at a cost to bee health.

Honey bees make propolis from a wide variety of resins depending on the diversity of resinous plants available in a particular environment (Crane, 1990). Like stingless bees, honey bees make choices between resinous plants that co-occur in the same environment by collecting from some plants, but not others (Leonhardt and Blüthgen, 2009; Wilson et al., 2013; Chapter 2). Plant resins are generally complex mixtures of phenolic and isoprenoid compounds that function in plant defense against pathogens and pests (Langenheim, 2003), though resin composition and antimicrobial activity can be highly variable between plants in different regions or even among species in the same genus (Lindenfelser, 1967; Bastos, et al., 2008; Wilson et al. 2013; Chapter 2).

Given the diversity of climate and potential botanical sources of resin in the U.S. (Crane, 1990; Wollenweber and Buchmann, 1999), it is probable that some regions will produce propolis with relatively greater or lesser antimicrobial activities. The antimicrobial diversity of U.S. propolis, as well as how that antimicrobial diversity might be biologically relevant to bees, is unknown. The goal of our current work is to characterize and compare the composition and antimicrobial activity of propolis from across the U.S. using LC-MS-based metabolomic methods and assays against biologically relevant bee pathogens. The composition and antimicrobial activity of U.S. propolis is generally understudied, and our aim is ultimately to better understand how variations in the botanical landscape may affect honey bee health. In addition, these

experiments should give insight into the diversity of bee preferred resinous plants in the U.S. and guide future studies focusing on the therapeutic potential of propolis for bees and isolation of useful antimicrobial compounds from U.S. propolis.

Materials and Methods

Collection of propolis samples

Propolis samples were collected from 12 locations in the U.S. representing distinct botanical regions from the desert southwest, to the cold-temperate north, to the coastal-temperate southeast (Chaska, MN; Baton Rouge, LA; Ithaca, NY; Jamestown, ND; Lincoln, NE; Raleigh, NC; Wakinsville, GA; Tucson, AZ; Aspen, CO; Vacaville, CA; Beaumont, TX; Fallon, NV) in 2009-2010 using commercial propolis traps (Mann Lake Ltd, cat. # HD-370). Propolis trapped from six different colonies per location was collected and pooled to create averaged regional samples, except in NV where 10 colonies were trapped. NV and MN propolis was re-sampled from five and one colonies, respectively, in 2012 for further analysis to confirm the reproducibility of antimicrobial activity. NV propolis was re-sampled ~20 miles from the original site and MN propolis was re-sampled ~30 miles from the original site. All samples were stored in sealed glass jars at -20°C.

Extraction

Frozen propolis samples were ground to a fine powder in a coffee grinder and 0.1 g of powdered material was extracted with 2 mL 70% ethanol by vortexing three times for 30 sec intervals. To remove any insoluble materials, the resulting extracts were chilled to precipitate extracted wax and centrifuged at 20,000 x *g* for 10 min. at 4°C with the resulting supernatants recovered. Extract concentration was measured by residue weight after solvent evaporation using vacuum centrifugation.

Analytical analysis

Propolis extracts were standardized to 1.5 mg/mL in 10% acetonitrile in water for metabolomic fingerprinting analysis by reversed-phase C_{18} liquid chromatography

[Thermo-Fisher (San Jose, CA) Acella LC system equipped with a Waters BEH C_{18} 1.0 x 100 mm, 1.8 µm particle size column; flow rate: 0.13 mL/min, column temperature: 35°C] coupled to Fourier transform mass spectrometry at 15,000 resolution [Thermo-Fisher Orbitrap XL, electro-spray ionization, negative ion mode (LC-FTMS)]. The Genedata Expressionist for Mass Spectrometry software package (http://www.genedata.com/products/expressionist/mass-spectrometry.html) was used to discover spectral peaks in the raw MS data in the form of mass/retention time pairs (peaks) based on Gaussian peak shape, absolute intensity, charge, and isotopic pattern. Multiply charged peaks and singlet peaks without an isotopic distribution were filtered out of the dataset. Comparative metabolite analyses were performed by performing pair-wise comparisons of the mass/retention time pair lists produced for each of the 12 regional propolis samples.

Identification of pinocembrin, phenylethyl caffeate, and pinobanksin-3-acetate in propolis samples using external standards was performed by re-analysis of propolis extracts using more powerful LC-FTMS conditions [Thermo-Fisher Dionex UltiMate 3000 LC system equipped with an Agilent (Santa Clara, CA) XDB-C_{18} 2.1 x 100 mm, 1.8 µm particle size column; flow rate: 0.4 mL/min, column temperature: 40°C, interfaced to a Thermo-Fisher Q-Exactive hybrid quadrupole orbitrap mass spectrometer at 17,500 resolution in negative ion mode]. Standard compounds were purchased from Sigma-Aldrich (pinocembrin, phenylethyl caffeate) or Angene Chemical (pinobanksin-3-aceate; Hong Kong).

Bacterial inhibition assay

Dilutions of 70% ethanol extracts were added to 96 well microplates and dried to a solvent-free residue under a stream of nitrogen gas. The *P. larvae* reference strain (NRRL #B-2605, ATCC 9545, LMG 9820 – ERIC type I) (de Graaf et al. 2006; Genersch et al. 2006), was obtained from the USDA Agricultural Research Service culture collection (http://nrrl.ncaur.usda.gov/). Liquid cultures of *P. larvae* were grown overnight with shaking at 37°C in brain/heart infusion broth (BHI, Difco) fortified with 1 mg/L thiamine and diluted 1:100 with fresh BHI in each microplate well. Total bacterial

growth was measured as the optical density of the well solution at 600 nm (OD_{600}) after 6 hrs of shaking and incubation at 37°C, which is midway through the growth phase of *P. larvae* in our conditions, with the OD_{600} of each well at 0 hrs subtracted as background. Relative bacterial growth for treated wells was calculated as the percent growth of untreated negative controls, with the absolute growth of untreated controls ~0.13 AU at t = 0 and ~0.60 at t= 6 hrs. Doses were 8, 10, 20, 30, 50, 60, 75, 100, 125, and 175 µg/mL of propolis extract, with 8 replicate wells per treatment. For comparison, the inhibitory activities of compounds previously reported as active against *P. larvae* (pinocembrin, pinobanksin-3-acetate, and phenyl caffeate) (Bilikova et al. 2012) were evaluated as above.

We also characterized the growth inhibition of tylosin (Sigma-Aldrich), an antibiotic approved for use against *P. larvae* infection, in our system using the same procedure as above, except dilutions were made in BHI and thus not dried to a residue before the addition of the liquid culture, as tylosin is very water soluble. The concentration range of tylosin used was 0.01 µg/mL to 10 µg/mL.

IC_{50} values were determined by growth curve analysis in SigmaPlot 10 (Systat Software Inc, Chicago, IL) by fitting the Hillslope equation to the sigmodial inhibition curves. Statistical significance between IC_{50} values was determined pair-wise between those samples that could be fit. 95% confidence intervals for the differences between IC_{50} values were calculated as $CI = z \pm [1.96 \times (\sqrt{x^2 + y^2})]$ where x is the standard error of $IC_{50(1)}$, y is the standard error of $IC_{50(2)}$, and z is the difference between $IC_{50(1)}$ and $IC_{50(2)}$. If the confidence interval of the difference between $IC_{50(1)}$ and $IC_{50(2)}$ did not overlap with 0, then the difference between the two IC_{50} values was taken as significant.

Fungal inhibition assays

A. apis reference strains were obtained from the ARS Entopathogenic Fungal Culture Collection (http://www.ars.usda.gov/is/np/systematics/fungibact.htm) [USDA #7405 (ATCC MYA-4450, mating type +) and USDA #7406 (ATCC MYA-4451, mating type -)]. Fungi were grown and mated on MY-20 media with spores isolated according to standard methods (Jensen et al., 2013). *A. apis* susceptibility to propolis was tested by

two microplate assays, one to evaluate spore germination and one to evaluate vegetative growth.

Microplates were pre-treated with propolis as above, except susceptibility tests were only performed with the 100 µg/mL dose. Spore germination was evaluated by inoculating each microplate well containing 180 µL of liquid MY-20 media with 8.1 x 10^4 spores in 20 µL of sterile water. In contrast, vegetative growth was evaluated by fragmenting the hyphae of USDA #7406 *A. apis* with mild homogenization in sterile water and then inoculating microplate wells containing 175 µL of liquid MY-20 media with 25 µL of the resulting suspension. We observed a long lag phase of ~50 hrs until growth started to occur, and total fungal growth was measured as above relative to untreated controls at 72 hrs, 96 hrs, and 7 days. These endpoints represented the growth phase, the stationary phase, and the distant stationary phases of growth. The Tukey-Kramer method for multiple comparisons was used to statistically test differences in growth from spores among propolis treatments instead of the confidence interval comparison used previously because we were able to compare measured values (mean OD_{600}) instead of inferred values (IC_{50} values) . Differences were considered significant at the 95% confidence level ($p \leq 0.05$).

Results

Diversity in propolis antimicrobial activity

In general, propolis samples inhibited the *in vitro* growth of *P. larvae* in a dose-dependent manner (Fig 1). Solvent controls showed no effects on bacterial growth from any lingering solvent after the drying of propolis residues. Regional propolis samples differentially inhibited bacterial growth, and statistical comparisons among 95% confidence intervals of differences between IC_{50} values supported four inhibitory groups (Table 1). Propolis samples from NV, CO, and TX were the most inhibitory ($IC_{50} \leq 50$ µg/mL, Table 1), while propolis samples from CA, AZ, NE, and ND were slightly less inhibitory ($IC_{50} \leq 80$ µg/mL, Table 1). Propolis from NC and GA showed some inhibitory activity, but did not completely inhibit *P. larvae* growth over the experimental

concentration range (Table 1, Fig 1). Propolis samples from MN, LA, and NY were relatively not very inhibitory (Table 1, Fig 1).

The antibiotic tylosin was much more inhibitory than any of the crude propolis extracts with IC_{50} = 0.255 ± 0.014 µg/mL; however, a direct comparison between tylosin and propolis is not possible because the composition of propolis extracts is complex, with the active compounds unknown and present in unknown concentrations. Surprisingly, pure pinocembrin, pinobanksin-3-acetate, and phenylethyl caffeate showed little to no growth inhibition in our assay with >90% growth of untreated controls at the highest tested concentration of 175 µg/mL.

Propolis from NV and MN were re-sampled to confirm their activities, as these were the most and among the least inhibitory samples, respectively. The new NV sample, sampled within 20 miles of the original site, showed similar activity to the original sample (IC_{50} = 52.8 ±4.8 µg/mL, Table 1); however, the new MN sample, sampled within 40 miles of the original site, showed increased activity (Estimated IC_{50} = 120 µg/mL, Table 1), but did not completely inhibit bacterial growth in the experimental concentration range.

There appeared to be some between colony variation in growth inhibition of *P. larvae* for NY propolis when propolis samples from different colonies were assayed independently (S1), which suggests that there may be similar variation among colonies from other regions, though sample pooling averaged out these differences in the regional comparison (Fig 1, Table 1).

Propolis samples were evaluated for an additional, bee-relevant antimicrobial activity by testing for the *in vitro* inhibition of *A. apis* growth. Regional propolis samples were assayed for activity at a propolis concentration of 100 µg/mL, a concentration at which we would expect ≥ 90% growth inhibition of *P. larvae* in most of the active samples. Seven of the 12 samples (CO, TX, CA, ND, NE, NV, and AZ) completely inhibited *A. apis* growth from spores at this dose (S2). Propolis samples also generally inhibited *A. apis* growth from direct inoculation of hyphal material (S3), though there was much more variability in growth compared to the spore inoculum.

Metabolomic analysis of propolis samples

Base-peak chromatograms of regional propolis samples are compared in Fig 2. 2,188 unique peaks were detected among the 12 propolis extracts, but only a subset of the total peaks appeared in any given sample (Fig 2, Table 2). The number of exclusive and pair-wise co-occurring peaks among all propolis extracts is shown in Table 2. Propolis samples with the highest activity against *P. larvae* that also showed high activity against *A. apis* (CO, TX, and NV) all shared a relatively large number of peaks, even though these samples came from botanically distinct regions. Although some samples had a relatively large number of exclusive peaks (CO, AZ, NC) there did not appear to be any relationship between the number of exclusive peaks and antimicrobial activity.

External standard analysis showed that pinocembrin, pinobanksin-3-acetate, and phenylethyl caffeate were present in all propolis extracts. The relative peak areas of these compounds in each sample did not correlate well ($R^2 < 0.5$) to the corresponding IC_{50} values, supporting a weak influence of these compounds on our observed antimicrobial activity against *P. larvae*.

Discussion

Our study supports that there is wide variation in the antimicrobial activities among propolis collected in the U.S., and that some regions harbor bee-preferred plant resins with greater or lesser antimicrobial activities that could be more or less beneficial to bees. Propolis samples were chemically distinct among different regions (Fig 2, Table 2), and differently inhibited the growth of *P. larvae* (Fig 1, Table 1) and *A. apis* (S2). Microbial inhibition ranged from good activity for a crude mixture to relatively non-inhibitory against both *P. larvae* and *A. apis* (Fig 1, Table 1, S2). Furthermore, the observed differences in anti-*P. larvae* activity were not due to previously reported compounds, as they were not active within our experimental concentration range, were present in all propolis samples, and the relative amounts of these compounds in the propolis samples were poorly correlated with bacterial growth inhibition.

The mechanistic benefit that bees derive from resins, assuming that greater antimicrobial activity is a benefit, is complex to consider. For example, propolis from

MN (second sampling) was not a good inhibitor of *P. larvae* growth, but was a good inhibitor of *A. apis* growth (Fig 1, S2). While many propolis samples were decent inhibitors of both *P. larvae* and *A. apis*, propolis from MN (second sampling) and NY displayed differential activity between pathogens in that they were more active against the fungal pathogen compared to the bacterial pathogen (Fig 1, S2). Even propolis samples that did not ultimately prevent full *A. apis* growth were able to slow growth through the early assay time points (S2). Honey bees are parasitized by at least 31 diverse organisms (Evans and Schwarz, 2011), and it is mostly unknown what role propolis may play against them. Evidence does suggests that there are specific functional relationships between resin collection and fungal pathogens (Chapusiat et al., 2007; Simone-Finstrom et al., 2013), but we do not know how pathogens are exposed to active resin compounds in colonies. Presumably, interaction with inhibitory compounds occurs via direct contact with resins, contact with resin compounds adhering to insect cuticles, or the volatilization of active resin compounds. *P. larvae* and *A. apis* are pathogenic through the ingestion of spores (Aronstein and Murray, 2010; Genersch, 2010) and, although our data show that propolis can prevention of fungal growth from spores, more research is needed determine the mechanism of propolis inhibition as bees are not known to consume propolis.

Metabolomic methods, including automated peak detection, made more information accessible from spectral data for comparisons without requiring the identification of specific compounds, which can be a very lengthy and expensive process. Not only were we able to determine *if* the samples were different from one another, but we were also able to determine overall *how* or *how much* the samples differed from one another. In general, U.S. propolis samples appeared to have a number of conserved, high intensity peaks (M-H ions: 271 m/z, 285 m/z, and 313 m/z among others), though each sample also contained a unique pattern of shared and exclusive peaks (Fig 2, Table 1). The most antimicrobial propolis samples shared a relatively large number of peaks (Table 1), but it is impossible to tell from these data if that resulted in their similar antimicrobial activities. Peak differences among samples are likely due to different resinous plants available in the various regions, though little exact information on the botanical sources

of propolis in the U.S. is available (Wollenwebber and Buchmann, 1997; Crane 1990; Wilson et al. 2013). The high number of peaks unique to propolis from Tucson, AZ (223 peaks), Raleigh, NC (101 peaks), and Aspen, CO (93 peaks) suggests that these regions have divergent botanical sources not shared across the entire sample group.

The purpose of this study was not to determine if propolis from 'here' is better than propolis from 'there', nor was it an attempt to find a 'cure' for bee disease, though clearly some regions produce propolis that is more antimicrobial than others (Fig 1, S2). Our study has uncovered potentially biologically relevant variations in both the antimicrobial activities among regional propolis samples and between pathogens, regional differences in propolis composition, and regions that likely harbor unique, bee-preferred resinous plants. In addition, this work provides a platform from which to rationally select U.S. propolis samples for the isolation of useful antimicrobial compounds. This work furthers our understanding of how the botanical landscape might influence bee health, but perhaps our greatest challenge in the future is to determine how active resin compounds interact with pathogens in the hive.

Acknowledgements

Thanks to C. Foster (Hidden Valley Honey, Reno, NV), T. Seely (Cornell University, Ithaca, NY), R. Oliver (Scientific Beekeeping, Grass Valley, CA), P. Limbach (Western Colorado Honey, Silt, CO), W. Klett (Jamestown, ND), S. Cobey (University of California-Davis, Davis, CA), N. Euliss (USGS, Jamestown, ND), D. Tarpy (North Carolina State University, Raleigh, NC), J. Villa (USDA-ARS, Baton Rouge, LA), M. Ellis (University of Nebraska, Lincoln, NE), D. Sammataro (USDA, Tuscon, AZ), and the University of Georgia Bee Lab for contributing propolis samples; D. Brinkman (University of Minnesota) for help developing and performing antimicrobial assays; M. P. Cohen (Principal Statistician, American Institutes for Research, Washington D.C.) for statistical help and advice; S. Cohen (Center for Regulatory Research, Roseville, MN) for help developing the fungal assays; G. Gardner, F. Gleason, and A. Hegeman (University of Minnesota) for reading and commenting on the manuscript; the Boswell/Cohen/Gardner/Hegeman and Spivak lab groups for useful

discussions; and the NSF Plant Genome Research Program grants IOS-0923960 and IOS-1238812, NSF grants IOS-0717530, IOS-1256992, and IOS-1256992, the Charles J. Brand Endowment, the UMN College of Food, Agriculture, and Natural Resource Sciences, the National Honey Board, the Minnesota Agricultural Experiment Station, and the Gordon and Margret Bailey Endowment for funding.

Figure 1 (a, b) – Inhibition of *P. larvae* growth by propolis extracts. Bacterial growth (y-axis) was measured as a percent optical density (OD_{600}) relative to untreated controls. There were 8 replicate wells per treatment for all propolis samples.

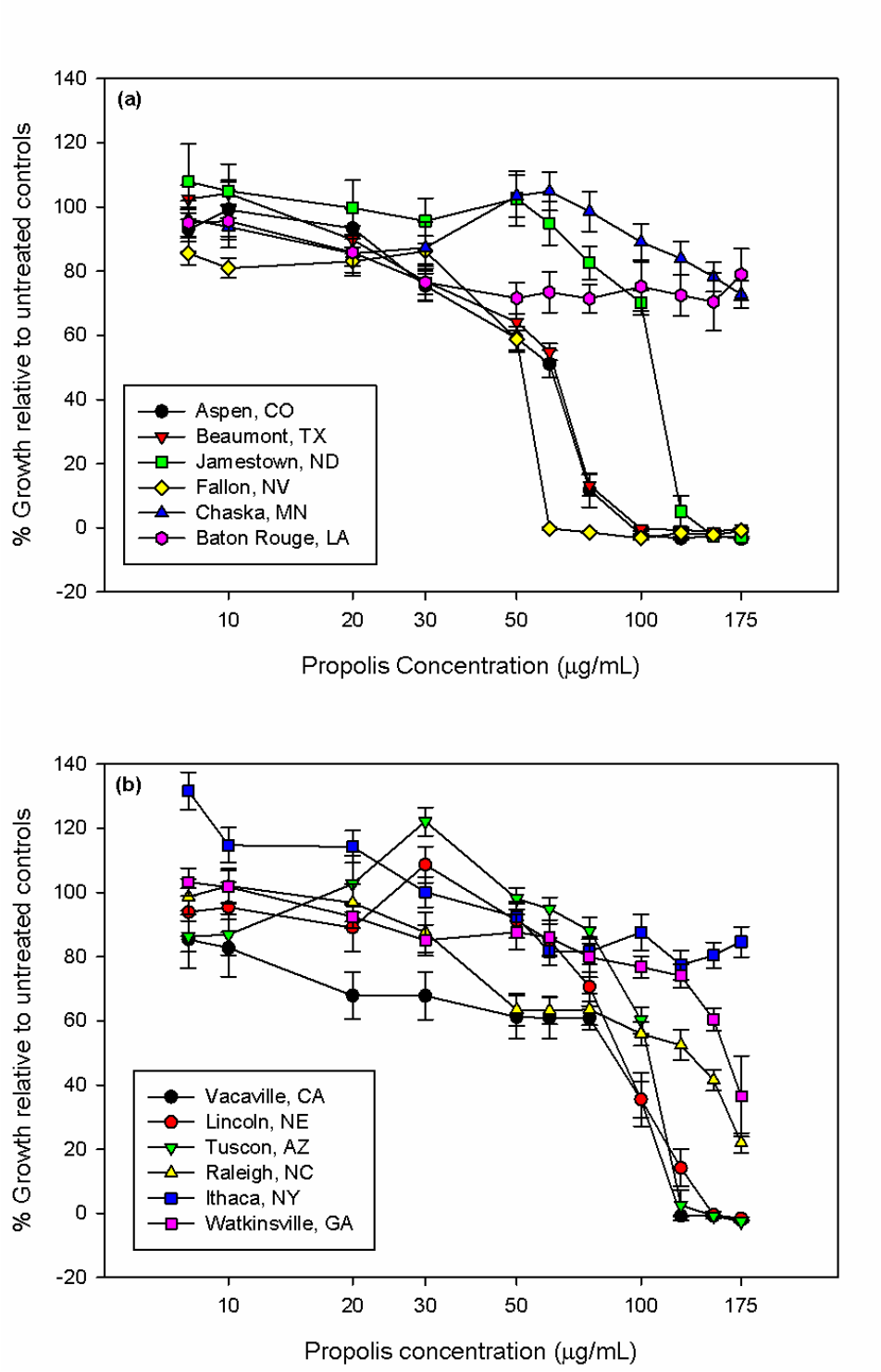
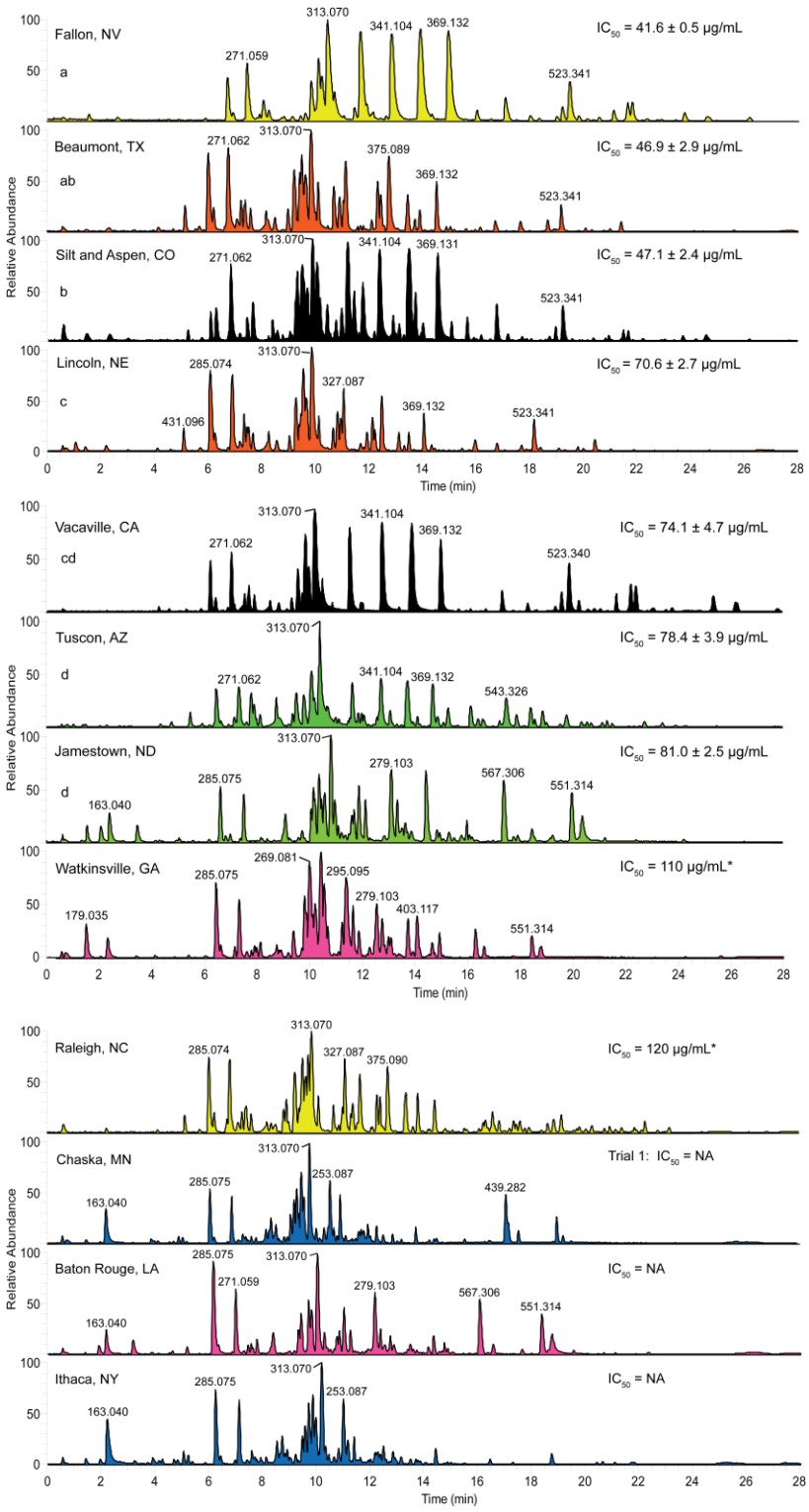
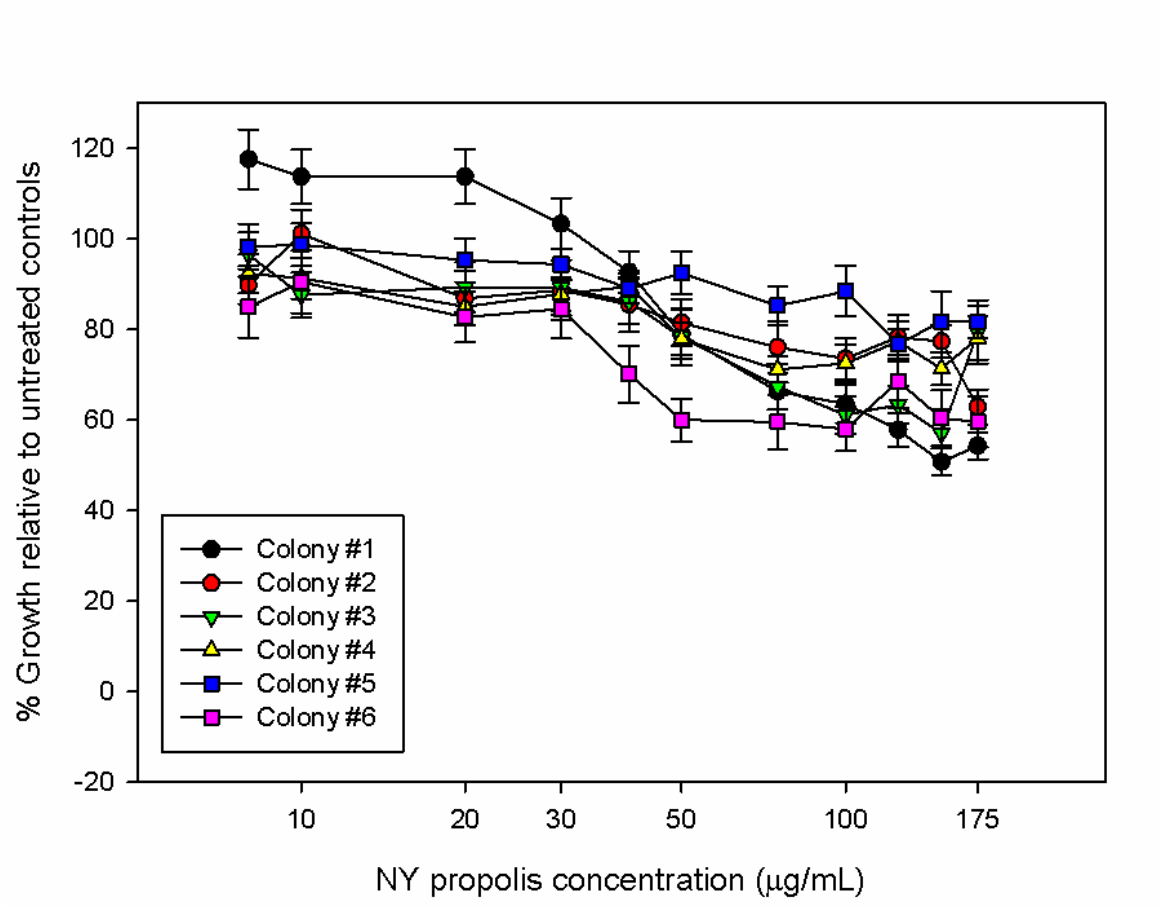


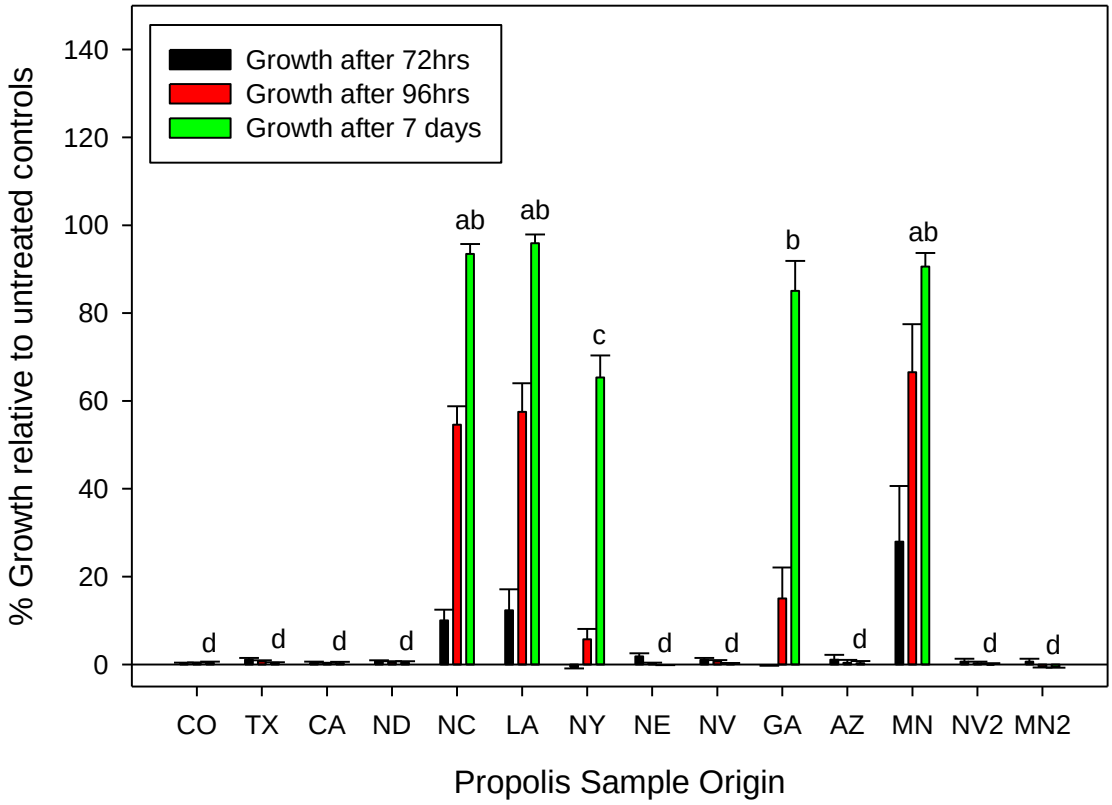
Figure 2 (a, b, c) – Base-peak, negative ion LC-FTMS chromatograms of propolis samples. Samples were standardized to 1.5 mg/mL for analysis. Peaks are annotated with the corresponding mass of the base-peak, and colors are coordinated with Fig 1. Mass accuracy was 2-5ppm. IC_{50} values calculated from growth curves (Fig 1, Table 1), is indicated for each sample. Statistical grouping, based on non-overlapping 95% confidence intervals for the IC_{50} values, is noted below the sample name. IC_{50} values annotated with (*) indicate that the value was estimated because the corresponding growth curve could not be fit. IC_{50} values of NA indicate that the sample had low or no activity over the experimental concentration range.


Fallon, NV
a
313.070
341.104
369.132
271.059
523.341
IC50 = 41.6 ± 0.5 µg/mL
Beaumont, TX
ab
271.062
313.070
375.089
369.132
523.341
IC50 = 46.9 ± 2.9 µg/mL
Silt and Aspen, CO
b
271.062
313.070
341.104
369.131
523.341
IC50 = 47.1 ± 2.4 µg/mL
Lincoln, NE
c
285.074
313.070
327.087
431.096
369.132
523.341
IC50 = 70.6 ± 2.7 µg/mL
Relative Abundance
Time (min)
Vacaville, CA
cd
313.070
341.104
369.132
271.062
523.340
IC50 = 74.1 ± 4.7 µg/mL
Tuscon, AZ
d
313.070
271.062
341.104
369.132
543.326
IC50 = 78.4 ± 3.9 µg/mL
Jamestown, ND
d
313.070
285.075
279.103
567.306
551.314
163.040
IC50 = 81.0 ± 2.5 µg/mL
Watkinsville, GA
269.081
285.075
295.095
279.103
403.117
179.035
551.314
IC50 = 110 µg/mL*
Relative Abundance
Time (min)
Raleigh, NC
313.070
285.074
327.087
375.090
IC50 = 120 µg/mL*
Chaska, MN
313.070
285.075
253.087
163.040
439.282
Trial 1: IC50 = NA
Baton Rouge, LA
285.075
313.070
271.059
279.103
567.306
551.314
163.040
IC50 = NA
Ithaca, NY
285.075
313.070
253.087
163.040
IC50 = NA
Relative Abundance
Time (min)

S1 – Colony variation in *P. larvae* inhibition by Ithaca, NY propolis. Bacterial growth (y-axis) was measured as a percent optical density (OD_{600}) relative to untreated controls. There were 8 replicate wells per treatment for propolis from all six colonies.



S2 – Inhibition of *A. apis* growth from spores by propolis. Fungal growth (y-axis) was measured as a percent optical density (OD_{600}) relative to untreated controls. Samples were evaluated at 100 µg/mL, which was a very inhibitory treatment in many of the *P. larvae* trials. NV2 and MN2 propolis denote the second sampling trial from Fallon, NV and St. Paul, MN, respectively. N = 8 replicate wells for CO, TX, CA, ND, NC, LA, and NY propolis. N = 5 replicate wells for NE, NV, AZ, GA, and MN propolis. N = 3 replicate wells for NV2 and MN2 propolis. Propolis treatments were grouped statistically after seven days using the Tukey-Kramer method for multiple comparisons at the 95% confidence level ($p \leq 0.05$). Grouping 'a' was not significantly different from untreated controls.



S3 – Inhibition of *A. apis* hyphal growth by propolis. Fungal growth (y-axis) was measured as a percent optical density (OD_{600}) relative to untreated controls. Samples were screened at 100 µg/mL, which was a very inhibitory treatment in many of the *P. larvae* trials. NV2 and MN2 propolis denote the second sampling trial from Fallon, NV and St. Paul, MN, respectively. N = 8 replicate wells for all propolis samples.

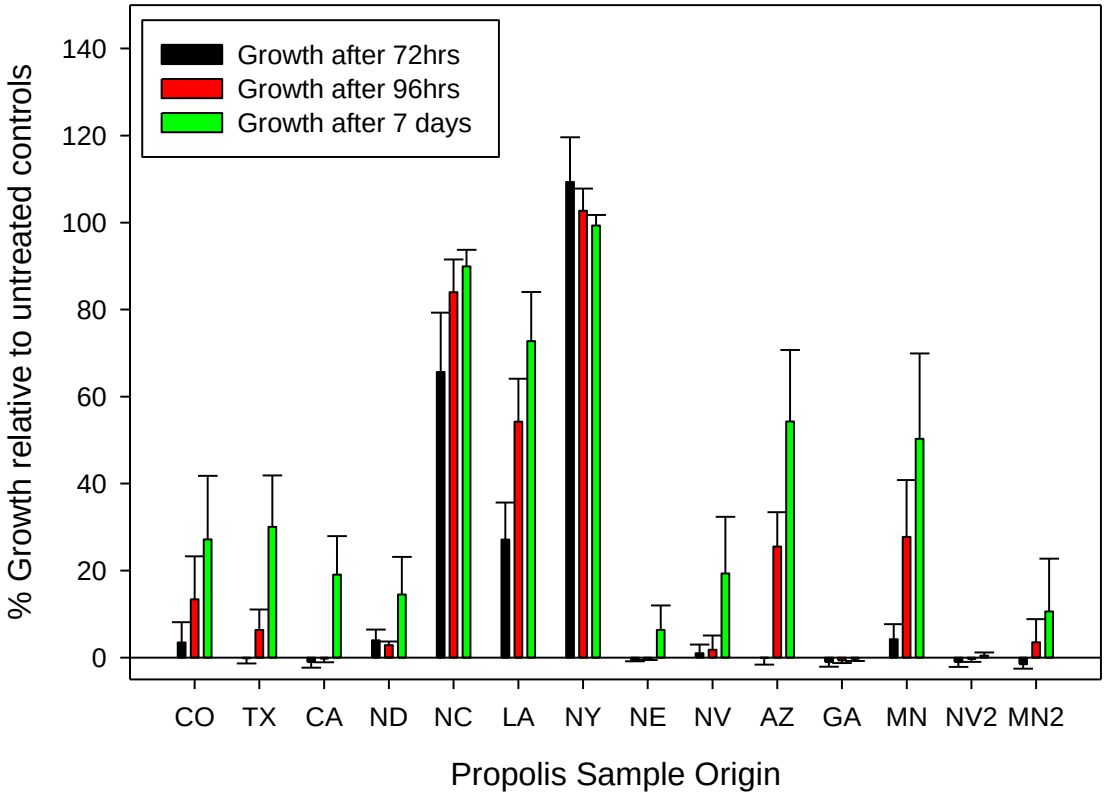


Table 1 – IC_{50} values of propolis samples inhibiting *P. larvae* calculated from dose-response curves. Values were calculated by fitting data points with a four-parameter curve in SigmaPlot 10. GoF indicates the "goodness of fit" of the four-parameter curve on the data points from each sample. (*) indicates the value was estimated because the corresponding growth curve could not be fit. IC_{50} values of NA indicate that the sample had low or no activity over the experimental concentration range. [(2)] indicates that these were part of the second trial, as described in the text. Statistical grouping, based on non-overlapping 95% confidence intervals for the IC_{50} values is noted following IC_{50} values.

Location	IC_{50} value (µg/mL)	GoF
Fallon, NV	41.6 ± 0.5 (a)	0.99
Beaumont, TX	46.9 ± 2.9 (ab)	0.97
Aspen, CO	47.1 ± 2.4 (b)	0.98
Lincoln, NE	70.6 ± 2.7 (c)	0.99
Vacaville, CA	74.1 ± 4.7 (cd)	0.96
Tucson, AZ	78.4 ± 3.9 (d)	0.96
Jamestown, ND	81.0 ± 2.5 (d)	0.99
Watkinsville, GA	110*	---
Raleigh, NC	120*	---
Chaska, MN	NA	---
Baton Rouge, LA	NA	---
Ithaca, NY	NA	---
Fallon, NV[(2)]	52.8 ± 4.8	0.96
St. Paul, MN[(2)]	120*	---

Table 2 – Compositional comparison of propolis extracts. Pair-wise comparison of peaks (single charged, negative ion mode) found in LC-FTMS chromatograms of propolis extracts from different regions in the U.S. The number of pair-wise co-occuring peaks between samples is shown in individual cells. Comparisons between a sample and itself (black cells) indicate the number of peaks that were exclusive to that sample. The total number of peaks detected in each sample is indicated in the last column. There were 2,148 unique peaks in the entire dataset.

	NV	TX	CO	NE	CA	AZ	ND	NC	LA	NY	GA	MN	Peak #
NV	18	371	435	104	291	211	224	304	165	175	212	175	560
TX		3	435	216	317	229	256	364	190	196	239	198	601
CO			93	211	294	224	246	340	185	191	232	197	870
NE				5	177	202	216	196	181	181	208	185	493
CA					9	182	189	265	155	153	179	165	474
AZ						223	229	207	174	180	218	187	948
ND							16	219	190	196	239	198	729
NC								101	170	173	204	180	672
LA									3	173	180	172	651
NY										9	190	181	759
GA											24	192	819
MN												47	756

Chapter 4
Isolation of flavanone-3-alkyl esters in western U.S. propolis active against bee pathogens

Summary

Honey bees, *Apis mellifera,* collect antimicrobial plant resins from their environments and use them as building materials in their nests. The presence of resin in the hive has been shown to have an array of positive physiological effects on bees, including specific roles against at least one bee pathogen. In Chapter 3 it was shown that propolis from Fallon, NV was particularly active against the bacterial bee pathogen *Paenibacillus larvae* as well as the fungal bee pathogen *Ascopheara apis* compared to propolis from other regions in the U.S. Using bioassay-guided fractionation against *P. larvae,* 11 structurally related flavanones from Fallon, NV propolis were isolated. This is the first reported NMR characterization of six of these isolated compounds, which had been previously characterized by MS alone one. One compound, pinobanksin-3-octanoate, has not been previously reported. Six of the isolated compounds were flavanone-3-alkyl esters which showed very high *in vitro* activity against both *P. larvae* and *A. apis*. There appeared to be both a positive and a negative relationship between the length of the alkyl ester and inhibition of *P. larvae* and *A. apis*, respectively. Two of the active compounds were shown to be major contributors to the anti-*P. larvae* activity previously reported for U.S. propolis samples in Chapter 3. Analysis of resin extracts from *Populus spp.* resin used in Chapter 2 indicated that poplar is a botanical source of the isolated active compounds.

Introduction

'Propolis' is the apicultural term for plants resins collected by honey bees, *Apis mellifera*, which are used as a building material in their nests. Enrichment of managed hives with propolis to simulate natural nesting behavior prevents chronic up-regulation of individual immune function compared to non-enriched hives, and bees bred to

accumulate more propolis tend to live longer and have healthier brood (Nicodemo et al., 2013; Simone et al., 2009). All of these outcomes can positively impact productivity, which makes resin collection an important aspect of bee biology for beekeepers. Propolis also has specific activity against microbial bee pathogens *in vitro*, including the bacteria *Paenibacillus larvae,* the causative agent of American foulbrood disease in bees, and the fungus *Ascospherea apis*, the filamentous ascomycete that causes chalkbrood disease in bees (Bastos et al., 2008; Lindenfelser, 1967; Chapter 3). It has also been shown that colonies challenged with *A. apis* respond by increasing resin collection, while artificial resin enrichment of the colony can prevent *A. apis* infection (Simone-Finstrom and Spivak, 2012). Similarly, it has been shown that the presence of conifer resins in the nests of wood ants, *Formica paralugubris*, another Hymenopteran insect, increases the survival of both adults and larvae when challenged with entopathogenic bacteria and fungi (Chapuisat et al., 2007).

It is thought that the benefits of resin to bees are driven by their antimicrobial properties, though bees in different regions collect a diversity of resins that differ in both activity and composition (Crane, 1990; Bankova, 2006, 2005; Wilson et al., 2013). It is unclear why bees collect some resins, but not others, when they are able to choose among several different resinous plants in their environment (Wilson et al., 2013). Resins are chemically complex and variable mixtures of phenolic and isoprenoid compounds, secreted by plants to protect against predators and pathogens (Langenheim, 2003), and propolis tends to have a general antimicrobial activity against gram-positive bacteria (Burdock, 1998; Kujumgiev et al., 1999); however, activities against fungi and gram-negative bacteria tend to be more variable by region (Burdock, 1998; Kujumgiev et al., 1999). In addition to the ability to directly inhibit microbial growth, propolis has been reported to increase the susceptibility of gram-positive (*Bacillus subtilus*) and gram-negative (*Escherichia coli*) bacteria to traditional antibiotics 1.2 to 1.75-fold, even when propolis treatment by itself has seemingly no effect on bacterial growth (Mirzoeva et al., 1997).

Lindenfelser (1967) conducted a comprehensive screen for antimicrobial activity in 15 different propolis samples against 25 different species of bacteria (including: *Mycobacterium spp., Pseudomonas spp., Xanthomonas spp.,* and *Bacillus spp.*) and 20 different species of fungi (including: *Aspergillus spp., Trichophyton spp.,* and *Claviceps purpurea*). That study found that at least one of the 15 propolis samples tested at a dosage of 100 µg/mL inhibited each individual microbe (Lindenfelser, 1967). Of the 45 pathogens tested, only *P. larvae* was inhibited by all 15 propolis samples at this dosage (Lindenfelser, 1967). In our previous evaluation of propolis activity against *P. larvae*, we observed IC_{50} values (the concentration at which 50% of growth is inhibited) between 41 µg/mL and 120 µg/mL for propolis sampled from 12 different geographic regions in the U.S. (Chapter 3). Nine of these propolis samples also completely inhibited *A. apis* at a 100 µg/mL dose (Chapter 3).

The antimicrobial compounds in propolis from temperate regions are generally thought to be flavonoids or organic acid esters derived from various species of poplar (*Populus spp.*) (Bankova, 2006, 2005b). The common flavonoids apigenin, chrysin, galangin, quercetin, naringenin, isosakuranetin, and pinocembrin have also all been found in European propolis and all have documented antimicrobial activities (Marcucci, 1995; Cushnie and Lamb, 2005). Pinocembrin and isosakuranetin treatment both inhibit *Staphylococcus aureus* growth with a minimum inhibitory concentration (the lowest concentration of a substance which results in no apparent growth – MIC) = 50 µg/mL, while also inhibiting *Candida albicans* growth with MIC = 25 µg/mL and 50 µg/mL, respectively (Rojas et al., 1992). Chrysin inhibits *C. albicans* growth with MIC = 25 µg/mL and also *Pseudomonas aeruginosa* growth with MIC = 31.25 µg/mL (Rojas et al., 1992). Quercetin and naringenin were found to reduce motility (linear velocity) of *Bacillus subtilis* by 5 and 2-fold, respectively, at a 20 µM dose, and motility is very important to the virulence of this bacterial species (Mirzoeva et al., 1997). Quercetin and apigenin were reported to inhibit bacterial DNA gyrase *in vitro* with high activity (Maximum quercetin activity occurred at 240 µM, while apigenin IC_{50} = 233 µM) (Ohemeng et al., 1993; Plaper et al., 2003), though antibacterial activity against *E. coli*

was weak without first disrupting the cytoplasmic membrane with polymyxin B nonapeptide. A 12 hr incubation with 50 µg/mL galangin was reported to reduce *S. aureus* growth by 60%, presumably through cytoplasmic membrane damage signaled by potassium leakage (Cushnie and Lamb, 2005b). Several phenolic compounds have been isolated from Bulgarian propolis (thought to be derived mainly from *Populus nigra* resin) with reported activity against *P. larvae* including pinocembrin, pinobanksin-3-acetate, and phenylethyl caffeate; however, these compounds were not responsible for the observed activity in U.S. propolis samples I studied (Chapter 3). Thus, it remains unknown what specific compounds are responsible for the antimicrobial activity of propolis from different regions in the U.S.

In Chapter 3, propolis from Fallon, NV was shown to have the highest activity among propolis samples evaluated for activity against *P. larvae* and *A. apis* with IC_{50} = 41 µg/mL and MIC < 100 µg/mL, respectively. Determining the compounds responsible for this antimicrobial activity will directly link specific products of plant metabolism to the benefits bees derive from particular resins. If the botanical sources of these antimicrobial compounds can be determined, then beekeepers can rationally modify the environment around their apiaries to make more antimicrobial resins available to their bees and also "mass produce" these compounds in a non-destructive, environmentally friendly manner. Propolis in the U.S. may have a great deal of untapped potential as a source of useful antimicrobial compounds, and discovering specific antimicrobial compounds in propolis will benefit beekeepers by adding value to what is considered a secondary hive product. The goal for the present work is to use bioassay-guided fractionation to establish a causal relationship between specific antimicrobial compounds in propolis and their derived health benefit to honey bees.

Materials and Methods

Sampling

Propolis from Fallon, NV, is the same sample utilized in Chapter 3. Briefly, propolis was collected from nine managed colonies of *Apis mellifera* in the Fallon, NV, area in 2008 with commercial propolis traps and stored at -20 °C until use.

Extraction and liquid-liquid partitioning

80.65 g of powdered propolis was extracted with 1 L of 70% ethanol/water (Decon labs, 200 proof ethanol; reverse osmosis, de-ionized and glass distilled water) for 48 hours, with 40 mins/day of sonication, followed by two additional extraction steps with 500 mL of 70% ethanol, using the same procedure. These ethanolic extracts were pooled, concentrated by rotary evaporation, and taken up in methanol (Sigma, Chromosolv Plus grade). The pooled extract was diluted to 10% methanol in water and partitioned successively against equal volumes of hexanes (Fisher Scientific, HPLC grade) twice for 24 hrs each time. The remaining water/methanol layer was then partitioned successively against equal volumes of dichloromethane (DCM) (Sigma, Chromosolv grade) twice for 24 hrs. The resulting three partitions (hexanes, dichloromethane, and water) were concentrated by rotary evaporation, weighed, and assayed for biological activity against *P. larvae*.

Preparatory flash and column chromatography

The Reveleris flash chromatography system (Grace – Deerfield, IL) was used for separations when indicated in normal-phase on 40 g silica cartridges (Grace) using a chloroform-methanol gradient (25 mL/min – time = 0, A = 100%, B = 0%; time = 5 min, A = 90%, B = 10%, time = 10.5 min, A = 88%, B = 12%; time = 15.5 min, A = 0%, B = 100%, time = 18 min, A = 0%, B = 100%) using evaporative light-scattering detection to monitor eluting compounds. Sephadex LH-20 (170 g, GE Healthcare Life Sciences – Pittsburg, PA) separations were performed when indicated in an open column (3 cm x 33.75 cm) in reversed-phase using methanol as the elution solvent.

Preparatory high-performance liquid chromatography (HPLC)

Pure compounds were isolated from partially purified fractions when indicated by preparatory reversed-phase HPLC (Agilent 1200 Series Preparatory HPLC system, Agilent Technologies – Santa Clara, CA) on a Zorbax Eclipse XDB C_{18} PrepHT 21.2 x 250 mm, 7µm particle size column (Agilent) using a water-methanol gradient (20 mL/min - time = 0 min, A = 60%, B = 40%; time = 20 min, A = 5%, B = 95%; time = 25 min, A = 5%, B = 95%). Eluting compounds were monitored for absorbance at 254 nm and 320 nm.

Thin layer chromatography (TLC)

Fractions were combined, when noted, based on TLC observations by spotting samples on silica gel 60 (F-254) (Merck, 5719-2) and developing in a solvent of 9.5:0.5 chloroform:methanol with 0.2% glacial acetic acid. Plates were evaluated by long and short UV and then by sulfuric acid/vanillin staining (15 g vanillin, 250 mL ethanol, and 2.5 mL concentrated sulfuric acid) with heat gun application.

Mass spectrometry analysis

Liquid chromatography mass spectrometry was used to both monitor fractionation and to characterize purified compounds. Fractions were monitored, when indicated, using an Acuity single-quadurpole LC-MS system (Waters – Milford, MA) using a Zorbax Eclipse XDB C_{18} 2.1 x 100 mm, 1.8 µm particle size column (Agilent). The reversed-phase separation was conducted with a water (0.1% formic acid) to acetonitrile (0.1% formic acid) gradient (time = 0 min, A = 90%, B = 10%; time = 20 min, A = 10%, B = 90%; time = 22 min, A = 10%, B = 90%; time = 25 min, A = 90%, B = 10%) Accurate mass and fragmentation analyses of purified compounds were conducted using a G2 Synapt LC-TOF system (Waters – Milford, MA) and an AmaZon LC-Ion trap MS (Bruker – Billerica, MA), respectively, utilizing the same separation conditions as above.

Antimicrobial assays

P. larvae growth inhibition assay was performed as described in Wilson et al. (2013) and Chapter 3 to guide propolis fractionation. Briefly, dilutions of assayed substances were added to 96 well microplates and dried to a solvent-free residue under a stream of nitrogen gas. Overnight liquid cultures of *P. larvae* [ARS culture collection, also known as the Northern Regional Research Laboratory collection (NRRL), #B-2605] were diluted 1:100 in each microplate well. Total bacterial growth was measured as the optical density of the well solution at 600 nm (OD_{600}) after 6 hrs of shaking and incubation at 37°C during the exponential growth phase. OD_{600} of each well at 0 hrs was subtracted as background. Relative bacterial growth was calculated as the percent growth of untreated negative controls and all tests were performed in triplicate. Concentration ranged from 8 µg/mL to 175 µg/mL, however a single concentration of 50 µg/mL was used to evaluate growth inhibition when material was limiting. 50 µg/mL is approximately the IC_{50} value of the crude 70% ethanol extract of Fallon, NV propolis (Chapter 3). A concentration range of 0.02 µg/mL to 50 µg/mL was used to test the inhibition of pure compounds.

Assays against *A. apis* were performed in 96 well plates similar to the procedure in Chapter 3, and only purified compounds were tested for activity against this pathogen in the present work. Briefly, spores were isolated from mated *A. apis* reference strains [USDA #7405 (ATCC MYA-4450, mating type +) and USDA #7406 (ATCC MYA-4451, mating type -)] for use in the susceptibility assay. Microplates were pre-treated with compounds as above, and 180 µL of liquid MY-20 media in each well was inoculated with 1.2×10^6 spores in 20 µL of sterile water. Again, a concentration range of 0.02 µg/mL to 50 µg/mL was used to test the inhibition of pure compounds. Microplates were shaken and incubated at 31°C, and growth from spores started to occur after 48 hrs. Relative fungal growth was calculated as the percent growth of untreated negative controls, and all tests were performed in triplicate.

Nuclear magnetic resonance (NMR) analysis

[^{1}H]-NMR, COSY, HMBC, and HMQC experiments (Crews et al., 2010) for compounds **7**-**9** and **11** were preformed on Varian VI-400 MHz and VI-500 MHz NMR spectrometers. HMQC and [^{13}C]-NMR experiments for compounds **1**-**6** and **10** were performed on a Bruker Avance 700 MHz NMR spectrometer with a TXI cryoprobe and a Bruker Avance III NMR with a TCI Cryoprobe, respectively. All compounds were analyzed in [$^{2}H_4$]methanol.

Results and Discussion

Compound isolation

Propolis was subjected to bioassay-guided fractionation by testing against the bee bacterial pathogen *P. larvae*. Extraction of 80.65 g of Fallon, NV propolis with 70% ethanol and subsequent liquid-liquid partitioning yielded 0.78 g in the hexanes partition, 49.6 g in the DCM partition, and 2.52 g in the remaining aqueous partition. Total recovery was 52.9 g with most of the insoluble material left behind appearing to be beeswax. Only the dichloromethane partition showed activity comparable to the crude extract with MIC < 100 µg/mL, while the hexanes layer showed slight activity with MIC > 150 µg/mL, and the aqueous partition showed no activity. Only the DCM layer was evaluated for further study.

Flash chromatography of 10 g of the DCM partition on silica produced 48 fractions which were combined based on thin-layer chromatography (TLC) similarities to 13 fractions yielding between 5.2 mg and 820.7 mg. Total recovery was 2.14 g. Three flash chromatography fractions (#5 – 432.9 mg, #6 – 512.6 mg, and #7 – 820.7 mg) displayed better or equivalent activity against *P. larvae* compared to the crude extract with MIC = 20 µg/mL, 30 µg/mL, and 60 µg/mL respectively. These fractions were evaluated using LC-MS and were combined based on a substantial number of overlapping compounds and further fractionated on Sephadex LH-20, producing 23, 20 mL fractions. Fractions were combined based on TLC and LC-MS similarities to 11 fractions yielding between 42.1 mg and 412 mg with a total recovery of 1.701 g.

Fractions 3-5 (122.6 mg, 206.2 mg, 129.7 mg) showed substantial activity with MIC = 10 µg/mL, 30 µg/mL, and 60 µg/mL respectively and were carried on for further study.

Sephadex LH-20 fractions 3-5 yielded 11 apparent compound(s) upon purification by preparatory HPLC. The Sephadex LH-20 fraction origin, retention time, and yield of compounds **1-11** can be found in Table 1. Additional quantities of each compound were purified as required for biological testing using the same procedures.

Compound identification

UV-Vis spectral analysis of isolated compounds suggested that all were flavanones or dihydroflavanols based on their UV-Vis absorption characteristics (λ_{max} = ~280 nm with a shoulder at ~310 nm) (Table 2) (Markham, 1982). LC-MS and LC-MS-MS analysis further suggested the isolated compounds were flavonoids through appropriate elemental compositions based on accurate mass and the production of A ring fragments characteristic of flavanones and flavan-3-ols (Pinheiro and Justino, 2010) (Table 3). In addition, the MS-MS data showed distinct losses supporting the presence of various length alkyl esters in compounds **1-6** and **9-10** (Table 3). All [^{1}H]-NMR spectra showed characteristic flavonoid patterns and confirmed that compounds **1-11** contained unsubstituted B rings with the presence of two multiplets around 7.4 ppm and 7.5 ppm and di-substituted A rings with the presence of two doublets around 5.96 ppm and 5.94 ppm (Fig 1a, 1b) (Marbry et al. 1970; Markham, 1982). Compounds **8**, **9**, and **11** were all isolated as a white powder with [M-H]- = 269.0830 *m/z*, 313.0717 *m/z*, and 271.0627 *m/z* respectively (Table 3). These compounds were confirmed to be pinostrobin (**8**), pinobanksin-3-acetate (**9**), and pinobanksin (**11**) by [^{1}H]-NMR and LC-MS-MS with comparison to authentic standards (Fig 2a). These compounds have been previously isolated from European propolis (Marcucci, 1995).

The [^{1}H]-NMR spectra of compounds **1-7** and **10** were consistent with the pinobanksin backbone observed in authentic standards. The [^{1}H]-NMR spectrum of compound **7** (white powder, [M-H]- = 285.0776 *m/z*] also indicated a single methoxylation at 3.83 ppm, but further analysis was required to determine if this

modification occurred at the 3, 5, or 7 position because the difference between δ_H of methoxy protons at any of these positions is small (MetIDB flavonoid NMR database, http://metidb.org/home). The $A^{1,3+}$ ion of 166.8 *m/z* in the MS-MS spectrum of compound **7**, as opposed to 152.8 *m/z* apparent in the other compounds, confirmed that the methoxylation must be at either position 5 or 7 (Table 3). UV-Vis shift analysis with NaOAc confirmed a free 7-OH with the observation of a +37 nm red-shift in Band II, supporting that the methoxylation must be at the 5 position (Markham, 1982). A red-shift in Band II observed with $AlCl_3$ and subsequent degradation back to starting λ_{max} with the addition of HCl indicated that there was no interaction between a free 5-OH and the 4-ketone (Markham, 1982), directly supporting the presence of a 5-OH methoxylation. All these results are consistent with compound **7** being identified as 5-methoxy-pinobanksin (**7**) (Fig 2b), which was previously isolated and characterized by NMR from Bulgarian propolis (Bankova et al., 1983).

Based on comparisons of MS-MS and [^{1}H]-NMR spectra from the pinobanksin-3-acetate standard and published NMR and MS data, compounds **1**-**6** and **10** were suspected to be pinobanksin-3- alkyl esters with different alkyl chain lengths, which have been previously reported in different propolis samples from temperate Europe and Uruguay (Marcucci, 1995; Kumazawa et al., 2002; Falcão et al., 2010). The presence of increasingly large ester losses, but otherwise very similar LC-MS-MS fragmentation patterns, for compounds **1**-**6** and **10** supports that these compounds are a family of very similar compounds with different length side chains (Table 3); however, it was impossible to determine the ester attachment site or alkyl branching pattern from MS-MS alone. HMBC NMR experiments confirmed that for compounds **1**-**6** and **10** the carbonyl of the alkyl ester side chain (position 1", Fig 1a) was connected to position 3 of the pinobanksin backbone through the observation of ^2J coupling between 1" and 3. We sometimes observed diestereotopic protons at position 2" on the alkyl chain protons, which was likely due to chirality at position 3 of the pinobanksin backbone. For most of our compounds, the signals between these two protons at 2" were poorly resolved at our field strength and considered as one multiplet in Fig 1a and Fig 1b.

Compounds **1** and **2** (white powder, [M-H]- = 341.1072 *m/z*) fragmentation patterns were consistent with the attachment of four carbon alkyl chains (Table 3). These compounds were isolated as isobaric isomers that were not resolvable by C_{18}, C_8, or NH_2 reversed-phase chromatography; however, we were still able to deduce the alkyl chain structure of each compound by NMR. The [H^1]-NMR and COSY spectra of these compounds indicated the presence of both the n- and iso- forms of a four carbon alkyl ester chain through the observation of a triplet at 0.75 ppm (n-chain, 4", 3H) coupled independently of two doublets at 0.9 ppm and 1.03 ppm (iso-chain, 3a" and 3b", 3H and 3H) (Fig 1b). Surprisingly, this difference in branching also led to differences in δ_H at positions 2 and 3 (Fig 1b). This combination of NMR experiments confirmed these compounds as pinobanksin-3-butyrate (**1**) and pinobanksin-3-(2-methyl)-propanoate (**2**) (Fig 2c) in approximately a 1:2 ratio. These compounds have both been previously reported in propolis from Iran and Europe by MS (Marcucci, 1995; Mohammadzadeh et al., 2007), through this is the first report of their characterization by NMR (Fig 2b).

Compounds **3** and **4** (orange oil, [M-H]- =355.1182 *m/z*) displayed fragmentation patterns consistent with the attachment of five carbon alkyl chains (Table 3). Similar to compounds **1** and **2,** compounds **3** and **4** differed slightly in their alkyl chain branching and were not chromatographically resolvable, though their structures could be deduced by [H^1]-NMR and COSY. The strong triplet at 0.55 ppm (4", 3H) systematically connected to a strong doublet at 1ppm (2"-CH_3) independent of two doublets at 0.74 ppm and 0.72 ppm (iso-chain, 4a" and 4b", 3H and 3H) indicated the presence of both 2-methyl branching and the iso-chain, respectively (Fig 2b). Like compounds **1** and **2**, this difference in branching caused relative shifts in δ_H at positions 2 and 3 (Fig 2b). Taken together, these data confirm compounds **3** and **4** as pinobanksin-3-isopentenoate (**3**) and pinobanksin-3-(2-methyl)butyrate (**4**) (Fig 2c, 2b) in approximately a 1:3 ratio. Compound **4** has been previously reported in Uruguayan propolis and characterized by NMR (Kumazawa et al., 2002), and compound **3** has been reported in Iranian propolis by MS (Mohammadzadeh et al., 2007); however, this is the first characterization of compound **3** by NMR (Fig 2b).

Compounds **5** (orange oil, [M-H]- = 397.1688 *m/z*) and **6** (orange oil, [M-H]- = 369.1391 *m/z*) displayed fragmentation patterns consistent with much longer alkyl chains relative to compounds **1-4**, with eight and six carbons respectively (Table 3). The lack of any doublets and the presence of a single triplet at ~0.9 ppm (8" and 6" respectively, 3H) in the [H^1]-NMR spectra of both compounds indicated that both contained n-alkyl esters (Fig 2b). The large number of methylene protons in this linear format caused considerable secondary coupling at 400 Mz H^1-NMR, though connectivity could be determined through COSY and HMBC experiments. These data considered together confirmed compounds **5** and **6** as pinobanksin-3-octanoate (**5**) and pinobanksin-3-hexanoate (**6**) (Fig 2d). This is the first unambiguous identification of compound **5** (Table 3, Fig 1b, Fig 2d) and the first report of NMR characterization of compound **6** (Fig 2d), though compound **6** has been previously reported in European propolis by MS (Marcucci, 1995).

Compound **10** displayed a mass and fragmentation pattern consistent with the addition of a single methylene on the alkyl ester chain compared to pinobanksin-3-acetate (**9**) (Table 3). The [^{1}H]-NMR spectrum of **10** is also consistent with the single addition of a methylene with in the alkyl chain with the presence of a multiplet at 2.24 ppm (2", 2H) (Fig 1b) confirming **10** as pinobanksin-3-propanoate (**10**) (Fig 1c). Compound 10 has been previously reported in European propolis (Marcucci, 1995), though this is the first report of its characterization by NMR (Fig 1b)

Inhibition of P. larvae and A. apis

Pure compounds were preliminarily evaluated for activity against *P. larvae* by testing growth inhibition at 50 µg/mL, with compounds **1-6** showing substantially greater activity than crude extract (crude extract inhibits ~50% of *P. larvae* growth at this concentration). Compounds **1-2** inhibited *P. larvae* growth by 60% at 50 µg/mL, while compounds **3-6** inhibited 100% of growth at this concentration. Compounds **7-11** were less active than the crude extract, inhibiting between 0% and 35% of growth and were not characterized further. The dose-responsiveness of *P. larvae* was characterized for

compounds **1-6** (Fig 3), and IC_{50} values with corresponding compound purities by LC-DAD are displayed in Table 4. All compounds were significantly more inhibitory than the crude extract, with IC_{50} values ranging from 17 µM to 68 µM (Table 4). For comparison, IC_{50} = 0.3 µM for the antibiotic tylosin used to treat *P. larvae* in the field (Chapter 3). There also appeared to be a positive structure-activity relationship between longer alkyl esters and *P. larvae* inhibition, with compounds **5** and **6** being the most inhibitory (Table 4). This is supported by a report that artificially acylating other flavonoids at the 3 position with alkyl esters greatly increased their efficacy against *S. aureus*, with maximum activity occurring with the addition of C_8 and C_{10} alkyl esters (Stapleton et al., 2004). Re-analysis of the LC-MS data on regional propolis samples in Chapter 3 showed a strong correlation between the relative peak areas of compounds **5** and **6** with lower IC_{50} values, indicating that these compounds were the major contributors to anti-*P. larvae* activity among previously studied U.S. propolis samples in Chapter 3 (Table 4).

Compounds **1-6** were also very effective inhibitors of *A. apis* growth (Table 4). Interestingly, there appeared to be an inverse of the bacterial structure-activity relationship where shorter alkyl esters corresponded to increased *A. apis* inhibition, with compound **1-4** being the most inhibitory and compounds **5-6** being less inhibitory (Table 4). *A. apis* previously showed the ability to overcome initial growth inhibition by some propolis samples (e.g. inhibition by NY propolis in Chapter 3), and the same was true of inhibition by compounds **1-4** and **6** after 96 hrs compared to 65 hrs**,** although to a much lesser extent (data not shown). Also, *A. apis* was not susceptible to the common fungicide diazolidinyl urea in the concentration range tested. The high activity of compounds **1-4** and **6** make it feasible that they contributed to the persistent *A. apis* growth inhibition shown by NV propolis in Chapter 3 (100% growth inhibition at 100 µg/mL after 72 hrs, 96hrs, and 7 days).

Re-analysis of resin extracts from representative greenhouse grown North American *Populus spp.* (*Populus nigra, P. fremontii, P. angustifolia, P. trichocarpa, P. deltoides,* and *P. balsamifera*) from Chapter 2 by LC-MS revealed the presence of

compounds **1-6** in multiple species. Notably, all species contained compounds **5-6**, though compounds **1-4** could not be detected in resin extracts from *P. angustifolia*, *P. deltiodes*, and *P. balsamifera* (Table 5). Flavanone-3-alkyl esters have been previously reported in *Populus spp.* resin (English et al., 1992, 1991; Greenaway and Whatley, 1990), but the configuration of their alkyl chains, and perhaps the attachment point of the alkyl esters, have previously been ambiguous. These results strongly suggest that *Populus spp.* are a bee-preferred (Chapter 2) environmental source of biologically active compounds, though more detailed analysis is required to determine if some poplar resins, or the propolis produced from them, contain varying amounts of compounds **1-6**.

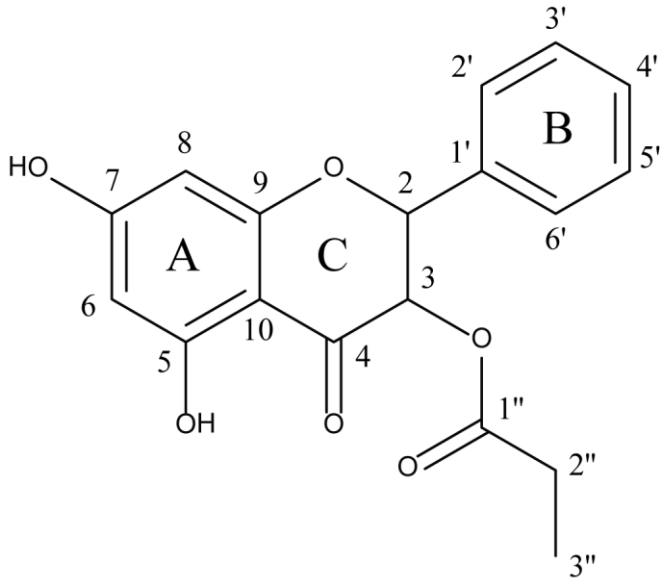
Conclusion

In this work bioassay-guided fractionation was used to isolate 11 structurally related compounds from Fallon, NV propolis. This is the first combined MS and NMR characterization of compounds **1-3**, **5-6**, and **10** (Fig 1a, 1b) and the first unambiguous identification of compound **5** (Table 3, Table 4, Fig 2b). Of these isolated compounds, half were determined to be flavanone-3-alkyl esters (compounds **1-6** and **10**) (Fig 2b, Fig 2c, Fig 2d). These flavanone-3-alkyl esters were also shown to have very high activity against the bee bacterial pathogen *P. larvae* and the bee fungal pathogen *A. apis* (Table 4). Synthesizing data from Chapter 3 with the inhibition data reported here revealed that compounds **5** and **6** were major contributors to the anti-*P. larvae* activity in propolis samples collected from different climatic regions in the U.S. (Chapter 3, Table 4). In addition, re-analysis of resin extracts from representative greenhouse grown North American poplars supported *Populus spp.* as the environmental source of these active compounds. Overall, this work provides a connection between specific compounds produced by *Populus spp.* and the benefits that bees can derive from foraging specifically from these plants in their environment.

Acknowledgements

I would like to thank Dr. Alison Pawlus (University of Minnesota) for her generous help during compound isolation and structure elucidation; Doug Brinkman (University of Minnesota) for maintaining bee pathogens and help collecting inhibition data; the USDA-ARS Culture Collections for providing reference strains of *P. larvae* and *A. apis*; Drs. Jerry Cohen, Gary Gardner, and Adrian Hegeman (University of Minnesota) for useful discussions and help interpreting data; Dr. Marla Spivak (University of Minnesota) for useful discussions and providing propolis via Randy Oliver (Scientific Beekeeping); the Boswell, Cohen, Gardner, Hegeman, and Spivak lab groups for useful discussions; and the NSF Plant Genome Research Program grants IOS-0923960 and IOS-1238812, NSF grants IOS-0717530, IOS-1256992, and IOS-1256992, the UMN College of Food, Agriculture, and Natural Resource Sciences, the National Honey Board, the Minnesota Agricultural Experiment Station, the Charles J. Brand Endowment, and the Gordon and Margaret Bailey Endowment for funding.

Figure 1a – NMR spectral data of isolated compounds for which there are no published reference spectra. Compounds **1-2** and **3-4** remained as isomeric mixtures after isolation, thus it was not possible to assign δ_C for these compounds. *J* values are reported in Hz. The numbering system of flavonoids is provided below. Refer to Fig 2c and 2d for structures.

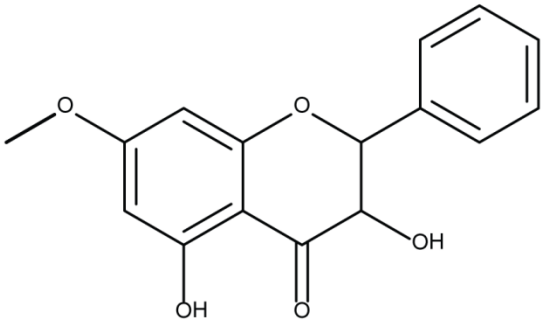


	Compound 1	Compound 2	Compound 3
	δ_H	δ_H	δ_H
2	5.4 d (*J* = 11.6)	5.42 d (*J* = 11.6)	5.39 d (*J* = 11.72)
3	5.8 d (*J* = 11.6)	5.86 d (*J* = 11.6)	5.86 d (*J* = 11.92)
6	5.94 d (*J* = 2)	5.94 d (*J* = 2)	5.95 d (*J* = 1.96)
8	5.96 d (*J* = 2)	5.96 d (*J* = 2)	5.97 d (*J* = 2.16)
2'	7.52 m	7.52 m	7.52 m
3'	7.41 m	7.41 m	7.4 m
4'	7.41 m	7.41 m	7.4 m
5'	7.41 m	7.41 m	7.4 m
6'	7.52 m	7.52 m	7.52 m
2"	2.22 m	2.48 dq (*J* = 7.2, 6.84)	2.11 m
3a"	1.47 dq (*J* = 7.5, 1.4)	0.9 d (*J* = 7.16)	1.86 m (J_{app} = 6.64)
3b"		1.03 d (*J* = 6.84)	
4a"	0.75 t (*J* = 7.2)		0.74 d (*J* = 6.64)
4b"			0.742 d (*J* = 6.64)

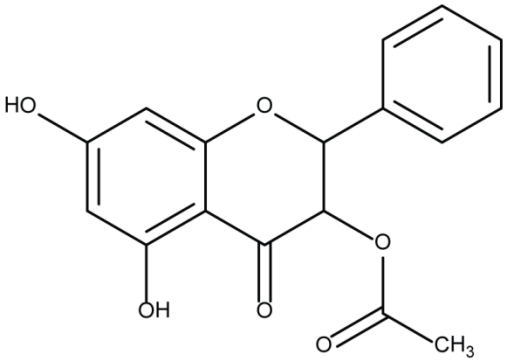
Figure 1b – NMR spectral data of isolated compounds for which there are no published reference spectra. *J* values are reported in Hz. Please refer to Fig 2c and 2d for structures, and also refer to text about diastereotopic protons at 2".

	Compound 5		**Compound 6**		**Compound 10**	
	δ_H	δ_c	δ_H	δ_c	δ_H	δ_H
2	5.42 d (*J* = 11.96)	72.12	5.42 d (*J* = 11.9)	74.42	5.36 d (*J* = 11.9)	74.03
3	5.87 d (*J* = 11.96)	81.21	5.87 d (*J* = 11.7)	83.7	5.82 d (*J* = 11.9)	83.03
4		194.49		194.46		194.32
5		163.6		165.62		165.34
6	5.94 d (*J* = 2.08)	95.44	5.96 d (*J* = 2.1)	99.04	5.96 d (*J* = 2.1)	97.98
7		168.2		171.79		170.22
8	5.95 d (*J* = 2.04)	96.54	5.98 d (*J* = 2.1)	101.46	6.01 d (*J* = 2.1)	99.08
9		163.6		165.62		165.34
10		100.55		103.15		103.33
1'		135.84		138.5		138.43
2'	7.54 m	127.46	7.54 m	130.2	7.47 m	130.01
3'	7.44 m	128.23	7.43 m	130.87	7.38 m	130.93
4'	7.44 m	129.04	7.43 m	131.78	7.38 m	131.81
5'	7.44 m	128.23	7.43 m	130.87	7.38 m	130.93
6'	7.54 m	127.46	7.54 m	130.2	7.47 m	130.01
1"		172.07		174.55		175.4
2"	2.26 m	33.17	2.26 m	33.25	2.24 m	28.99
3"	1.43 m	24.5	1.44 m	35.85	0.93 t (*J* = 7.6)	10.43
4"	1.1 m	28.35	1.09 m	26.95		
5"	1.21 m	28.57	1.21 m	33.3		
6"	1.31 m	22.24	0.85 t (*J* = 7.3)	24.5		
7"	1.31 m	22.24				
8"	0.89 t (*J* = 6.84)	13.01				

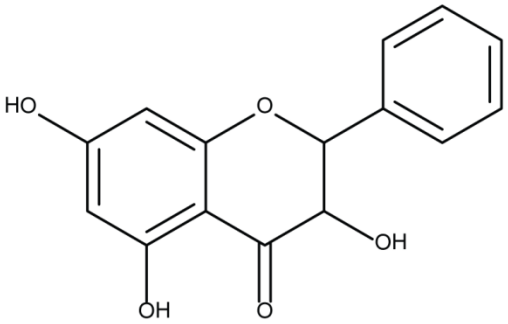
Figure 2a – Compounds isolated from NV propolis. The structure of compounds **8**, **9**, and **11** were confirmed by comparison to authentic standards using LC-MS-MS and ^{1}H-NMR. All have been previously reported in propolis.



Pinostrobin (8)

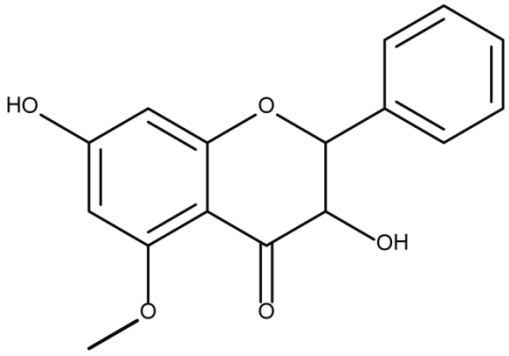


Pinobanksin-3-acetate (9)

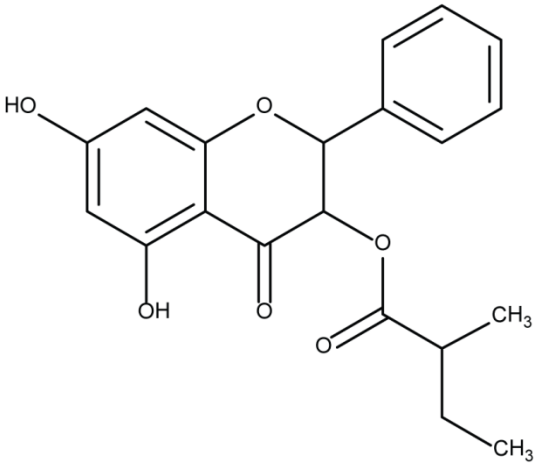


Pinobanksin (11)

Figure 2b – Compounds isolated from NV propolis. Compounds **2** and **4** were confirmed by LC-MS-MS fragmentation analysis and UV-Vis shift analysis (**2**) or comparison to documented H^1-NMR spectra (**4**). Both have been previously reported in propolis.

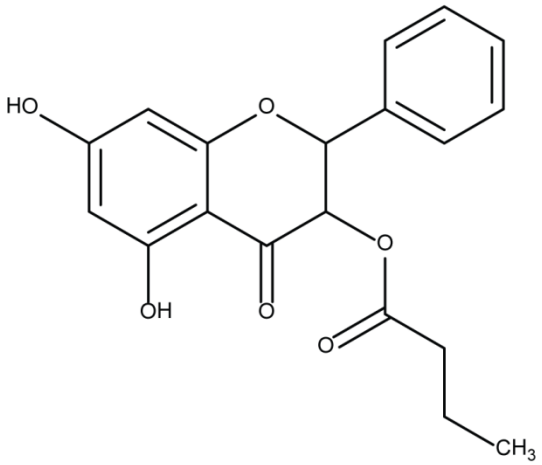


Pinobanksin-5-methyl-ether (2)

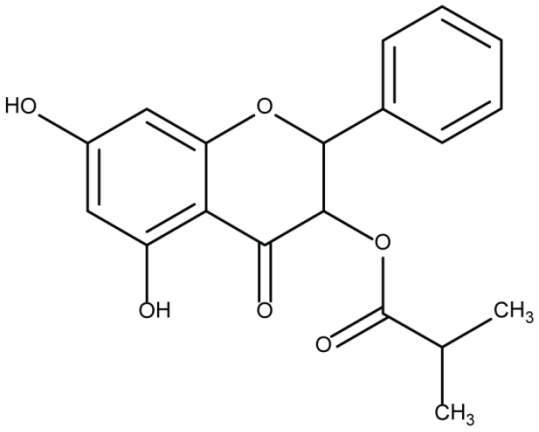


Pinobanksin-3-(2-methyl)-butyrate (4)

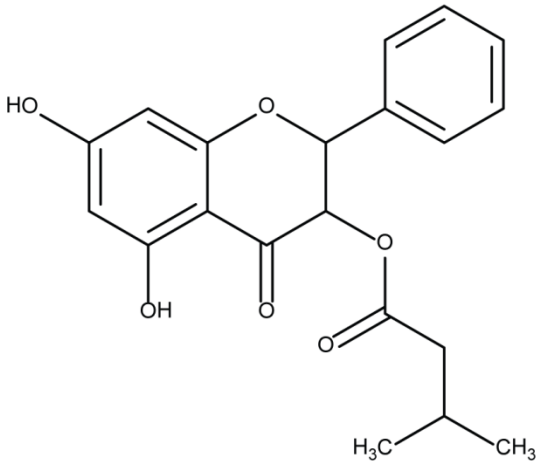
Figure 2c – Compounds isolated from NV propolis. Isolated short-chain flavanone-3-alkyl esters (compounds **1-3** and **10**) for which there are no published NMR reference spectra, but that have been reported to be in propolis through the use of MS fragment analysis alone.



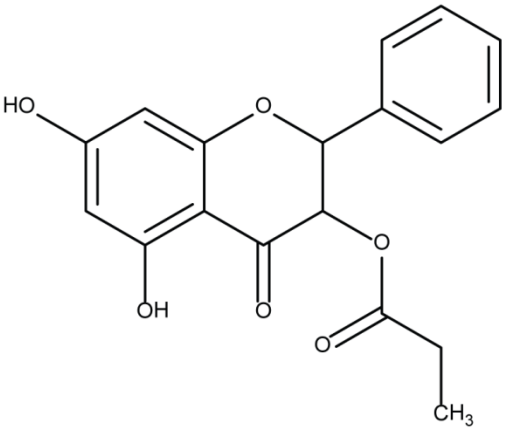
Pinobanksin-3-butyrate (1)



Pinobanksin-3-(2-methyl)-propanoate (2)

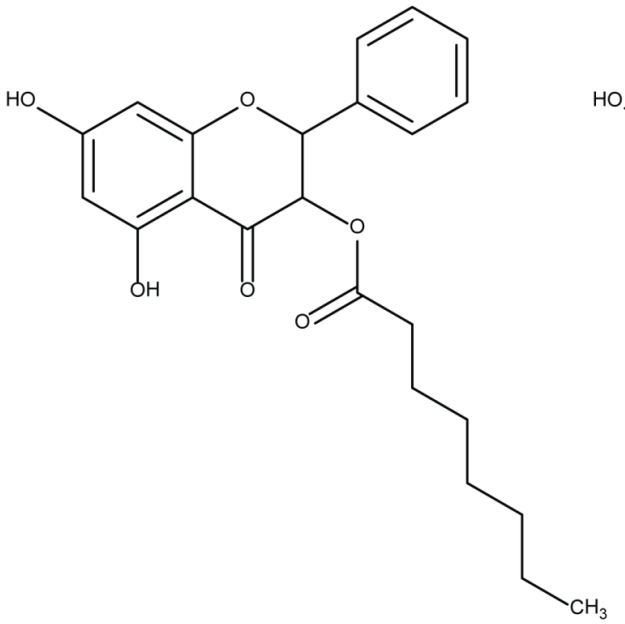


Pinobanksin-3-isopenanoate (3)

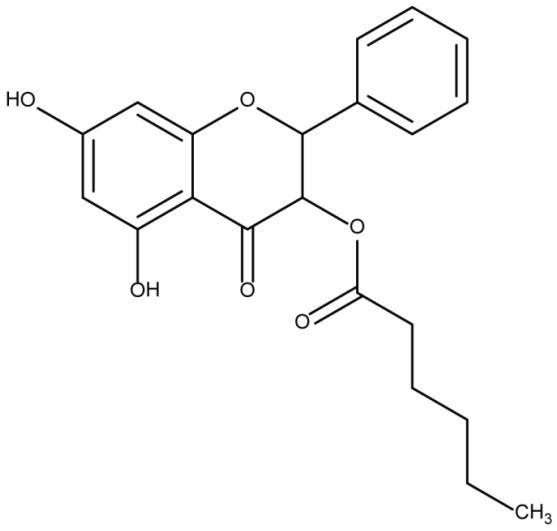


Pinobanksin-3-propanoate (10)

Figure 2d – Compounds isolated from NV propolis. Isolated long-chain flavanone-3-alkyl esters (compounds **5** and **6**) for which there are no published reference NMR spectra. Compound **6** had been previously reported to be in propolis through MS fragment analysis alone. Compound **5** has not been previously reported



Pinobanksin-3-octanoate (5)



Pinobanksin-3-hexanoate (6)

Figure 3 – Inhibition of *P. larvae* (a) and *A. apis* (b) by compounds 1-6. Compounds 1-2 and 3-4 could not be separated on the preparative scale, and so were tested as mixtures. N = 3 for all treatments.

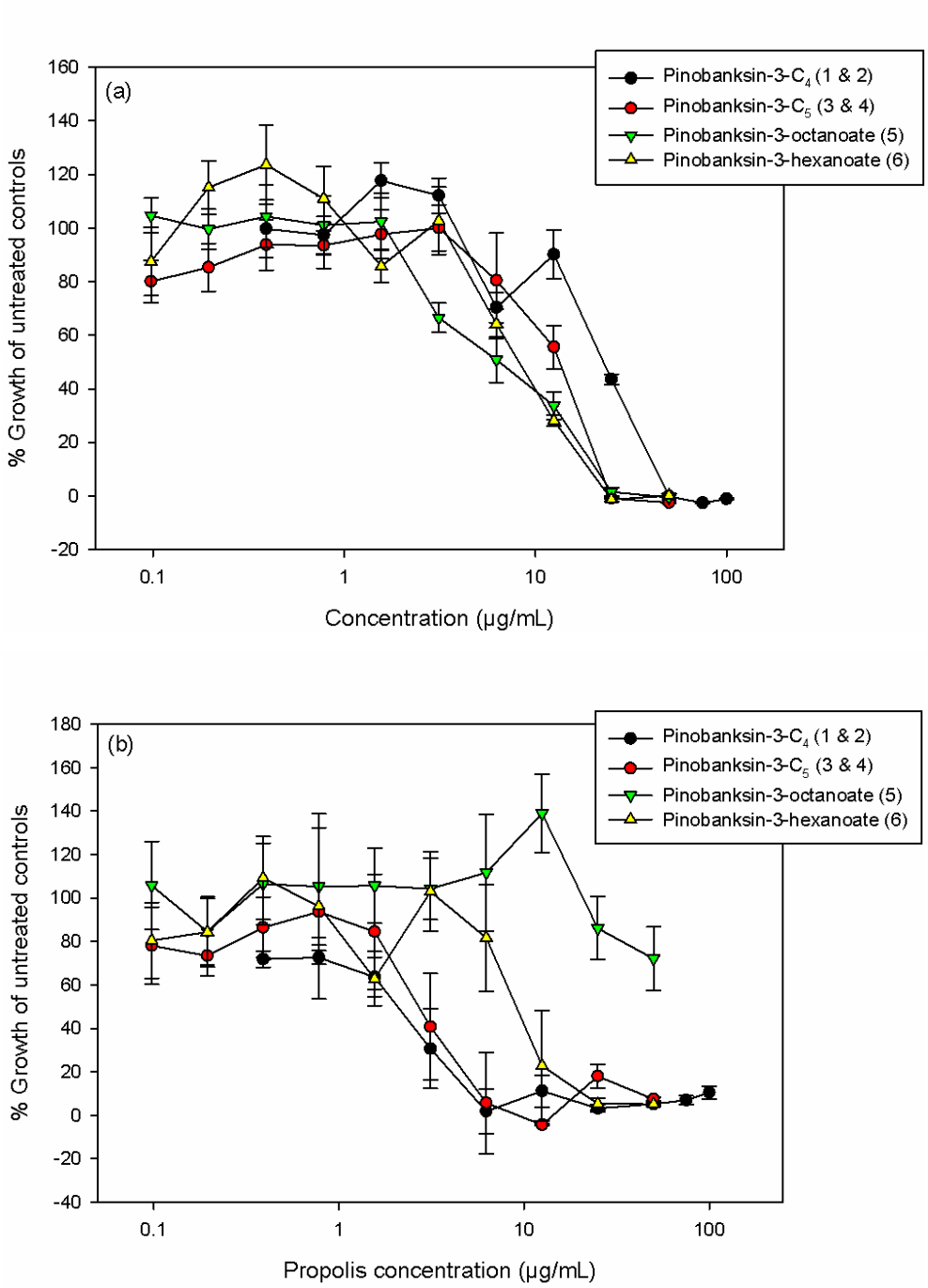
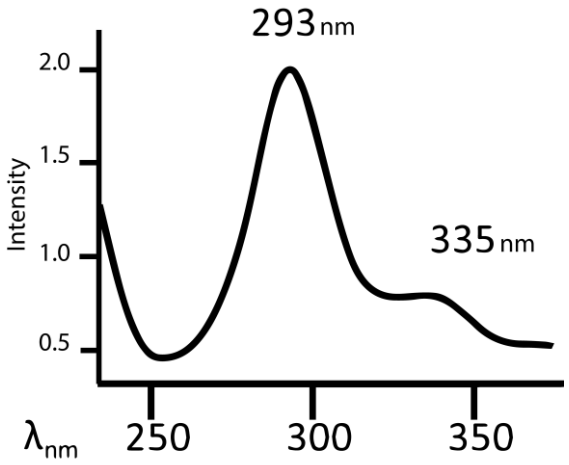


Table 1 – Sephadex fraction origin, preparative HPLC retention time, and yield of isolated compounds 1-11. HPLC conditions are detailed in the methods section. Compounds **1-2** and **3-4** were isolated as isomeric mixtures so yields given for compounds **1** and **3** include compounds **2** and **4**, respectively.

Compound	Origin	Retention time (min)	Yield (mg)
1	3	19	24.5
2	3	19	---
3	3	20.2	45.5
4	3	20.2	---
5	3	23	3.6
6	4	21.2	1.8
7	4	11	2
8	4	13.3	0.8
9	4	15.7	72.2
10	4	17.5	44
11	5	11.6	3.5

Table 2 – UV spectra of compounds 5-11. All compound displayed spectra characteristic of flavanones (Markman, 1982) and a sample spectrum from pinobanksin-3-acetate is shown. Spectra from pinobanksin-derived compounds were all extremely similar, though A-ring methylation on compounds **7** and **8** caused noticeable spectral shifts. Compounds **1-4** are not shown because pure spectra could not be obtained.



Compound	λ_{max} (nm)	$\lambda_{shoulder}$ (nm)
Pinobanksin-5-methyl ether (7)	288	315
Pinostrobin (8)	286	320
Pinobanksin (11)	293	335
Pinobanksin-3-acetate (9)	293	335
Pinobanksin-3-propanoate (10)	293	335
Pinobanksin-3-hexanoate (6)	294	335
Pinobanksin-3-octanoate (5)	293	335

Table 3 – Mass spectrometry data for compounds 1-3, 5, 6, and 10. '*m/z*' denotes the *m/z* of the $[M-H]^-$ ion observed with high mass accuracy as described in the methods section. Positive ion fragments >10% of the base-peak are shown with unit mass resolution.

Compound	Exact mass	*m/z*	Fragments	Fragment Annotation
1 & 2	342.1103	341.1072	343.0	$[M+H]^+$
			272.9	Partial alkyl ester loss - C_4H_7O
			254.9	Full alkyl ester loss - $C_4H_7O_2$
			226.9	Loss of CO
			152.8	Predicted $A^{1,3+}$ ion
3	356.1260	355.1182	357.0	$[M+H]^+$
			272.9	Partial alkyl ester loss - C_5H_9O
			254.9	Full alkyl ester loss - $C_5H_9O_2$
			226.9	Loss of CO
			152.8	Predicted $A^{1,3+}$ ion
5	398.1729	397.1688	399.5	$[M+H]^+$
			272.8	Partial alkyl ester loss - $C_8H_{15}O$
			254.9	Full alkyl ester loss - $C_8H_{15}O_2$
			226.8	Loss of CO
			152.8	Predicted $A^{1,3+}$ ion
6	370.1416	369.1391	371.1	$[M+H]^+$
			272.9	Partial alkyl ester loss - $C_6H_{11}O$
			254.9	Full alkyl ester loss - $C_6H_{11}O_2$
			226.8	Loss of CO
			152.8	Predicted $A^{1,3+}$ ion
10	328.0947	327.0872	329.0	$[M+H]^+$
			272.8	Partial alkyl ester loss - C_3H_5O
			254.8	Full alkyl ester loss - $C_3H_5O_2$
			226.8	Loss of CO
			152.8	Predicted $A^{1,3+}$ ion

Table 4 – Inhibition of *P.larvae* and *A. apis* growth by compounds 1-6. Compounds **1-2** and **3-4** could not be separated and so were assayed as mixtures. Reported sample purity was determined by LC-DAD scanning from λ_{210} to λ_{500}. IC_{50} values were calculated from growth curves shown in Fig 3. 'Goodness of fit' indicates how well the inhibition data could be fit by the Hillslope equation to calculate IC_{50} values. (*) indicates that the compound did not completely inhibit growth in the concentration range, thus no calculation of IC_{50} could be made. LC-MS data from the regional propolis samples in Chapter 3 were re-analyzed to determine if these compounds were contributing to overall bacterial inhibition in that sample set. The negative correlation between relative peak area of each compound(s) in a given propolis sample and its corresponding IC_{50} value is given in the last column (e.g. How true is it that increased peak area leads to lower IC_{50} values?). Higher R^2 values indicate stronger relationships.

Compound(s)	Purity	*P. larvae* IC_{50} (µM)	Goodness of fit (R^2)	Correlation with *P. larvae* inhibition of propolis extracts in Chapter 3 (R^2)	*A. apis* IC_{50} (µM)	Goodness of fit (R^2)
1-2	98%	68 ± 17	0.94	0.51	8 ± 0.5	0.98
3-4	98%	39 ± 4	0.97	0.33	8 ± 0.5	0.99
6	98%	22 ± 5	0.94	0.79	23 ± 2	0.99
5	82%	17 ± 4	0.98	0.88	*	---

Table 5 – Appearance of compounds 1-6 in *Populus spp.* resin. Resin extracts from representative greenhouse-grown *Populus spp.* from Chapter 2 were normalized to 1 mg/mL and re-analyzed by LC-MS at unit resolution. Isolated compounds were matched to their corresponding peaks in the resin samples by *m/z* and retention time. '+' denotes that the compound was present, while 'ND' denotes that this compound was 'not detected' in the sample.

Compound(s)	***P. nigra***	***P. fremontii***	***P. angustifolia***	***P. trichocarpa***	***P. deltoides***	***P. balsamifera***
1-2	+	+	ND	+	ND	ND
3-4	+	+	ND	+	ND	ND
5	+	+	+	+	+	+
6	+	+	+	+	+	+

Conclusion

The goal of this research has been to understand what resinous plants bees utilize and what benefits they might derive from these resins. In Chapter 2, I found that bees have many available resinous plants that are diverse in both composition and antimicrobial activity, and bees clearly favored some plants over others. Individual resin foraging bees were chemically tracked using metabolomics methods to determine that both *Populus deltoides* and *Populus balsamifera* are botanical sources of resin at the St. Paul campus apiaries. The metabolomic methods used in my analyses represent substantial improvements over previous analyses by increasing discrimination power between compositionally similar samples in a descriptive way, while also eliminating the need for deep biochemical knowledge of a resin to identify its botanical source. Surprisingly, *P. deltoides* and *P. balsamifera* resins were not the most antimicrobial resins available to bees within two miles of the apiary. Following the idea that trees from the genus *Populus* seem to be preferred resin sources, analyses were conducted to determine the compositional and antimicrobial differences among North American poplars. LC-MS based metabolomics indicated that *Populus spp.* were chemically distinct and that many spectral peaks were conserved to terminal taxonomic nodes in the *Populus* phylogeny. *Populus spp.* also differed in their ability to inhibit the *in vitro* growth of the bee bacterial pathogen *Paenibacillus larvae,* suggesting that all poplars may not be equally beneficial to bees.

In Chapter 3, I further explored the concept that plants in different regions may be more or less beneficial to bees. Although we do not fully understand how resin metabolites interact with microbes in honey bee nests, antimicrobial activity against bee pathogens is likely a benefit conferred to bee colonies by resins. In this research, I evaluated the ability of propolis samples from 12 climatically diverse regions in the United States for their ability to inhibit the growth of *P. larvae* and the fungal bee pathogen *Ascophaera apis*. Propolis samples displayed a range of activities, from relatively inactive to IC_{50} values < 50 µg/mL for *P. larvae* and MIC values < 100 µg/mL

for *A. apis*. Furthermore, the observed differences in anti-*P. larvae* activity were not due to previously reported compounds. This study uncovered potentially biologically relevant variations in the antimicrobial activities among regional propolis samples, but the specific compounds responsible for these activities remained unclear.

The goal of Chapter 4 was to isolate and characterize the compounds responsible for the observed antimicrobial activity against bee pathogens. Using bioassay guided fractionation against *P. larvae*, I isolated 11 structurally related compounds from Fallon, NV propolis (IC_{50} = 41 µg/mL against *P. larvae* and MIC < 100 µg/mL against *A. apis*). Half of these compounds were determined to be flavanone-3-alkyl esters with differing alkyl chains that had not been previously characterized by NMR. In addition, this was the first unambiguous identification of one compound, pinobanksin-3-octanoate. Several of the isolated flavanone-3-alkyl esters, including pinobanksin-3-octanoate, were very effective inhibitors of *P. larvae* and *A. apis* growth *in vitro* with IC_{50} values ranging from 7 µg/mL to 23 µg/mL against *P. larvae* and 3 µg/mL to 9 µg/mL against *A. apis*. Synthesizing data from Chapter 3 with the inhibition data reported here revealed that the two most active compounds were major contributors to the anti-*P. larvae* activity observed in propolis samples collected from different climatic regions in the U.S. In addition, re-analysis of resin extracts analyzed in Chapter 2 supported that *Populus spp.* are the environmental source of these very active compounds. Overall, this work provided a connection between specific antimicrobial compounds likely derived from *Populus spp.* and the benefits that bees can derive from foraging specifically from these plants in their environment.

Bibliography

Abraham, M.H., Ibrahim, A., Zissimos, A.M., 2004. Determination of sets of solute descriptors from chromatographic measurements. J. Chromatogr. A 1037, 29–47.

Aga, H., Shibuya, T., Sugimoto, T., 1994. Isolation and identification of antimicrobial compounds in Brazilian propolis. Biosci. Biotechnol. Biochem. 58, 945–946.

Alaux, C., Ducloz, F., Crauser, D., Le Conte, Y., 2010. Diet effects on honeybee immunocompetence. Biol. Lett. 6, 562–5.

Alfonsus, E.C., 1933. Some sources of propolis. Glean. Bee Cult. 61: 92-93

Allsopp, M.H., de Lange, W.J., Veldtman, R., 2008. Valuing insect pollination services with cost of replacement. PLoS One 3, DOI: 10.1371/journal.pone.0003128.

Amoros, M., Sauvager, F., Girre, L., Cormier, M., 1992. *In vitro* antiviral activity of propolis. Apidologie 23, 231–240.

Anderson, D., Munro, A., 1969. An analytical study of gum exudates from the genus *Araucaria* jussieu (gymnospermae). Carbohydr. Res. 11, 43–51.

Antúnez, K., Harriet, J., Gende, L., Maggi, M., Eguaras, M., Zunino, P., 2008. Efficacy of natural propolis extract in the control of American Foulbrood. Vet. Microbiol. 131, 324–31.

Armbruster, W., 1984. The role of resin in angiosperm pollination: ecological and chemical considerations. Am. J. Bot. 71, 1149–1160.

Aronstein, K. a, Murray, K.D., 2010. Chalkbrood disease in honey bees. J. Invertebr. Pathol. 103, S20–9.

Arrhenius, S.P. and Langenheim, J.H., 1983. Inhibitory effects of *Hymenaea* and *Copaifera* leaf resins on the leaf fungus, *Pestalotia subcuticularis*. Biochem. Sys. Ecol. 11(4): 361-366

Arteaga, S., Andrade-Cetto, A., Cárdenas, R., 2005. *Larrea tridentata* (Creosote bush), an abundant plant of Mexican and US-American deserts and its metabolite nordihydroguaiaretic acid. J. Ethnopharmacol. 98, 231–9.

Aslam, M., Choudhary, B., Uzair, M., Ijaz, A., 2013. Phytochemical and Ethno-Pharmacological Review of the Genus *Araucaria*–Review. Trop. J. Pharm. Res. 12, 651–659.

Atapattu, S.N., Poole, C.F., 2008. Solute descriptors for characterizing retention properties of open-tubular columns of different selectivity in gas chromatography at intermediate temperatures. J. Chromatogr. A 1195, 136–45.

Babić, S., Pašić, Š., Dukić, B., Ćoralić, A., 2011. Antifungal activity of propolis originated from Bosnia and Herzegovina. Veterinaria 60, 187–193.

Bankova, V., 2005a. Chemical diversity of propolis and the problem of standardization. J. Ethnopharmacol. 100, 114–7.

Bankova, V., 2005b. Recent trends and important developments in propolis research. Evid. Based. Complement. Alternat. Med. 2, 29–32.

Bankova, V., 2006. Plant sources of propolis: an update from a chemist's point of view. Nat. Prod. Commun. 1, 1023–1028.

Bankova, V., Boudourova-Krasteva, G., Popov, S., Sforcin, J., Cunha Funari, S., 1998. Seasonal variations of the chemical composition of Brazilian propolis. Apidologie 29, 361–367.

Bankova, V., Castro, S. de, Marcucci, M., 2000. Propolis: recent advances in chemistry and plant origin. Apidologie 31, 3–15.

Bankova, V., Popov, S., Marekov, N., 1983. A study on flavonoids of propolis. J. Nat. Prod. 46, 471–474.

Banskota, a H., Tezuka, Y., Adnyana, I.K., Ishii, E., Midorikawa, K., Matsushige, K., Kadota, S., 2001a. Hepatoprotective and anti-*Helicobacter pylori* activities of constituents from Brazilian propolis. Phytomedicine 8, 16–23.

Banskota, a H., Tezuka, Y., Kadota, S., 2001b. Recent progress in pharmacological research of propolis. Phytother. Res. 15, 561–71.

Barnes, B.B., Wilson, M.B., Carr, P.W., Vitha, M.F., Broeckling, C.D., Heuberger, A.L., Prenni, J., Janis, G.C., Corcoran, H., Snow, N.H., Chopra, S., Dhandapani, R., Tawfall, A., Sumner, L.W., Boswell, P.G., 2013. "Retention projection" enables reliable use of shared gas chromatographic retention data across laboratories, instruments, and methods. Anal. Chem. 85, 11650–7.

Bastos, E.M.A.F., Simone, M., Jorge, D.M., Soares, A.E.E., Spivak, M., 2008. *In vitro* study of the antimicrobial activity of Brazilian propolis against Paenibacillus larvae. J. Invertebr. Pathol. 97, 273–81.

Becerra, J.X., 2001. Interactions Between Chemical and Mechanical Defenses in the Plant Genus *Bursera* and Their Implications for Herbivores. Integr. Comp. Biol. 41, 865–876.

Bilikova, K. et al., 2012. New anti-*Paenibacillus larvae* substances purified from propolis. Apidologie, 44(3), 278-285

Boecking, O., Spivak, M., 1999. Behavioral defenses of honey bees against *Varroa jacobsoni* Oud. Apidologie 30, 141–158.

Boege, K., Marquis, R.J., 2005. Facing herbivory as you grow up: the ontogeny of resistance in plants. Trends Ecol. Evol. 20, 441–8.

Boswell, P.G., Carr, P.W., Cohen, J.D., Hegeman, A.D., 2012. Easy and accurate calculation of programmed temperature gas chromatographic retention times by back-calculation of temperature and hold-up time profiles. J. Chromatogr. A 1263, 179–88.

Boswell, P.G., Schellenberg, J.R., Carr, P.W., Cohen, J.D., Hegeman, A.D., 2011. A study on retention "projection" as a supplementary means for compound identification by liquid chromatography-mass spectrometry capable of predicting retention with different gradients, flow rates, and instruments. J. Chromatogr. A 1218, 6732–41.

Burdock, G. a, 1998. Review of the biological properties and toxicity of bee propolis (propolis). Food Chem. Toxicol. 36, 347–63.

Cane, J., 2002. Pollinating bees (Hymenoptera: Apiformes) of US alfalfa compared for rates of pod and seed set. J. Econ. Entomol. 95, 22–27.

Canto-Aguilar, M., Parra-Tabla, V., 2000. Importance of conserving alternative pollinators: assessing the pollination efficiency of the squash bee, *Peponapis limitaris* in *Cucurbita moschata* (Cucurbitaceae). J. Insect Conserv. 4, 203–210.

Chapuisat, M., Oppliger, A., Magliano, P., Christe, P., 2007. Wood ants use resin to protect themselves against pathogens. Proc. Biol. Sci. 274, 2013–7.

Christe, P., Oppliger, A., Bancalala, F., Castella, G., Chapuisat, M., 2002. Evidence for collective medication in ants. Ecol. Lett. 6, 19–22.

Corsano, S., Mincione, E., 1965. Triterpenoids from *Mangifera indica*: I. The structure of mangiferolic acid. Tetrahedron Lett. 28, 2377–2381.

Cowan, M., 1999. Plant products as antimicrobial agents. Clin. Microbiol. Rev. 12.

Cox-Foster, D.L., Conlan, S., Holmes, E.C., Palacios, G., Evans, J.D., Moran, N. a, Quan, P.-L., Briese, T., Hornig, M., Geiser, D.M., Martinson, V., vanEngelsdorp, D., Kalkstein, A.L., Drysdale, A., Hui, J., Zhai, J., Cui, L., Hutchison, S.K., Simons, J.F., Egholm, M., Pettis, J.S., Lipkin, W.I., 2007. A metagenomic survey of microbes in honey bee colony collapse disorder. Science 318, 283–7.

Crane, E., 1990. Bees and beekeeping: science, practice, and world resources. Comstock Pub. Associates.

Cremer, S., Armitage, S.A.O., Schmid-Hempel, P., 2007. Social immunity. Curr. Biol. 17, R693–702.

Crews, P., Rodríguez, J., Jaspars, M., 2010. Organic Structure Analysis. Oxford University Press.

Curtis, J., Lersten, N., 1974. Morphology, seasonal variation, and function of resin glands on buds and leaves of *Populus deltoides* (Salicaceae). Am. J. Bot. 61, 835–845.

Cushnie, T., Lamb, A., 2005. Detection of galangin-induced cytoplasmic membrane damage in *Staphylococcus aureus* by measuring potassium loss. J. Ethnopharmacol. 101, 243–8.

Cushnie, T.P.T., Lamb, A.J., 2005. Antimicrobial activity of flavonoids. Int. J. Antimicrob. Agents 26, 343–356.

Daugsch, A., Moraes, C.S., Fort, P., Park, Y.K., 2008. Brazilian red propolis--chemical composition and botanical origin. Evid. Based. Complement. Alternat. Med. 5, 435–41.

de Gaaf, K. et al., 2006. Identification of *Paenibacillus larvae* to the subspecies level: An obstacle for AFB diagnosis. J. Invert. Patho. 91, 115-123

Di Prisco, G., Cavaliere, V., Annoscia, D., Varricchio, P., Caprio, E., Nazzi, F., Gargiulo, G., Pennacchio, F., 2013. Neonicotinoid clothianidin adversely affects insect immunity and promotes replication of a viral pathogen in honey bees. Proc. Natl. Acad. Sci. U. S. A. 110, 18466–18471.

Drummond, F., Aronstein, K., Chen, J., Eitzer, B., Ellis, J., Evans, J., Ostiguy, N., Sheppard, W., Spivak, M., Visscher, K., 2013. The first trial of the stationary hive project: Abiotic site effects. http://www.beeccdcap.uga.edu/

Ellis, J.D., Hepburn, H.R., Ellis, a. M., Elzen, P.J., 2003. Social encapsulation of the small hive beetle (*Aethina tumida* Murray) by European honeybees (*Apis mellifera L.*). Insectes Soc. 50, 286–291.

English, S., Greenaway, W., Whatley, F., 1991. Analysis of phenolics of *Populus trichocarpa* bud exudate by GC-MS. Phytochemistry 30, 531–533.

English, S., Greenaway, W., Whatley, F., 1992. Analysis of phenolics in the bud exudates of *Populus deltoides*, *P. fremontii*, *P. sargentii* and *P. wislizenii* by GC-MS. Phytochemistry 31, 1255–1260.

Evans, J.D., 2003. Diverse origins of tetracycline resistance in the honey bee bacterial pathogen *Paenibacillus larvae*. J. Invertebr. Pathol. 83, 46–50.

Evans, J.D., Aronstein, K., Chen, Y.P., Hetru, C., Imler, J.-L., Jiang, H., Kanost, M., Thompson, G.J., Zou, Z., Hultmark, D., 2006. Immune pathways and defence mechanisms in honey bees *Apis mellifera*. Insect Mol. Biol. 15, 645–56.

Evans, J.D., Pettis, J.S., 2005. Colony-level impacts of immune responsiveness in honey bees, *Apis mellifera*. Evolution 59, 2270–4.

Evans, J.D., Schwarz, R.S., 2011. Bees brought to their knees: microbes affecting honey bee health. Trends Microbiol. 19, 614–20.

Evans, J.D.and Spivak, M., 2009. Socialized medicine: Individual and communal disease barriers in honey bees. J. Invertebr. Path. 103: S62-S72

Falcão, S.I., Vilas-Boas, M., Estevinho, L.M., Barros, C., Domingues, M.R.M., Cardoso, S.M., 2010. Phenolic characterization of Northeast Portuguese propolis: usual and unusual compounds. Anal. Bioanal. Chem. 396, 887–97.

Fearnly, J., 2001. Bee Propolis: Healing from the Hive. Souvenir Press, London UK

Gallai, N., Salles, J.-M., Settele, J., Vaissière, B.E., 2009. Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. Ecol. Econ. 68, 810–821.

Garedew, A., Schmolz, E., Schricker, B., Lamprecht, I., 2002. Microcalorimetric investigation of the action of propolis on *Varroa destructor* mites. Thermochim. Acta 382, 211–220.

Gekker, G., Hu, S., Spivak, M., Lokensgard, J.R., Peterson, P.K., 2005. Anti-HIV-1 activity of propolis in CD4(+) lymphocyte and microglial cell cultures. J. Ethnopharmacol. 102, 158–63.

Genersch, E., 2010. American Foulbrood in honeybees and its causative agent, *Paenibacillus larvae*. J. Invertebr. Pathol. 103 Suppl, S10–S19.

Genersch, E., et al., 2006. Reclassification of *Paenibacillus larvae* subsp. *pulvifaciens* and *Paenibacillus larvae* subsp. *larvae* as *Paenibacillus larvae* without subspecies differentiation. Int. J. Syst. Evol. 56, 501-511

Gershenzon, J., Dudareva, N., 2007. The function of terpene natural products in the natural world. Nat. Chem. Biol. 3, 408–14.

Ghisalberti, E.L., 1976. Propolis: a review. Bee World 60: 59-84

Ghosal, S., Biswas, K., Chattopadhyay, B., 1978. Differences in the chemical constituents of *Mangifera indica*, finfected with *Aspergillus niger* and *Fusarium monilformae*. Phytochemistry 17, 689–694.

Gilliam, M., Taber III, S.T., Lorenz, B., Prest, D., 1988. Factors affecting development of chalkbrood disease in colonies of honey bees, *Apis mellifera*, fed pollen contaminated with *Ascosphaera apis*. J. Invertebr. Pathol. 52, 314–325.

Goulson, D., Lye, G.C., Darvill, B., 2008. Decline and conservation of bumble bees. Annu. Rev. Entomol. 53, 191–208.

Greenaway, W, English, S., May, J., Whatley, F., 1991a. Analysis of phenolics of bud exudate of *Populus sieboldii* by GC-MS. Phytochemistry 30, 3005–3008.

Greenaway, W., English, S., May, J., Whatley, F.R., 1991. Chemotaxonomy of section Leuce poplars by GC-MS of bud exudate. Biochem. Syst. Ecol. 19, 507–518.

Greenaway, W, Gümüsdere, I., Whatley, F., 1991b. Analysis of phenolics of bud exudate of *Populus euphratica* by GC-MS. Phytochemistry 30, 1883–1885.

Greenaway, W., Scaysbrook, T., Whatley, F.R., 1987. The Analysis of Bud Exudate of *Populus euramericana*, and of Propolis, by Gas Chromatography--Mass Spectrometry. Proc. R. Soc. B Biol. Sci. 232, 249–272.

Greenway, W., Jobling, J., Scaysbrook, T., 1987. Composition of bud exudate of *Populus x interamericana* clones as a guide to clonal identification. Silvae Genetica 38(1): 28-32

Greenaway, W., Scaysbrook, T., Whatley, F.R., 1988. Phenolic analysis of bud exudate of *Populus laszocarpa* by GC-MS. Phytochemistry 27, 3513–3515.

Greenaway, W., Whatley, F., 1990. Analysis of phenolics of bud exudate of *Populus angustifolia* by GC-MS. Phytochemistry 29, 2551–2554.

Greenaway, W., Whatley, F., 1991. Analysis of phenolics of bud exudate of *Populus ciliata* by GC-MS. Phytochemistry 30, 1887–1889.

Greenleaf, S.S., Kremen, C., 2006. Wild bee species increase tomato production and respond differently to surrounding land use in Northern California. Biol. Conserv. 133, 81–87.

Hamzeh, M. and Dayanandan, S., 2004. Phylogeny of *Populus* (Salicaceae) based on nucleotide sequences of chloroplast trnT-trnF region and nuclear rDNA. Am. J. Bot. 91(9): 1398-1408

Higes, M., Martín-Hernández, R., Botías, C., Bailón, E.G., González-Porto, A. V, Barrios, L., Del Nozal, M.J., Bernal, J.L., Jiménez, J.J., Palencia, P.G., Meana, A., 2008. How natural infection by *Nosema ceranae* causes honeybee colony collapse. Environ. Microbiol. 10, 2659–69.

Hornitzky, M., 2001. Literature review of chalkbrood, a fungal disease of honey bees. Rural Industries Research and Development Corporation. RIRDC Publication No 01/150

Isidorov, V. A., Vinogorova, V.T., 2003. GC-MS analysis of compounds extracted from buds of *Populus balsamifera* and *Populus nigra*. Z. Naturforsch. C. 58, 355–60.

Johnson, R., 2010. Honey Bee Colony Collapse Disorder. Congr. Res. Serv. RL33938.

Johnson, R.M., Ellis, M.D., Mullin, C.A., Frazier, M., 2010. Pesticides and bee toxicity – USA. Apidologie 41: 312-331

Johnson, R.M., Dahlgren, L., Siegfried, B.D., Ellis, M.D., 2013. Acaricide, fungicide and drug interactions in honey bees (*Apis mellifera*). PLoS One 8, e54092.

Klein, A.-M., Vaissière, B.E., Cane, J.H., Steffan-Dewenter, I., Cunningham, S.A., Kremen, C., Tscharntke, T., 2007. Importance of pollinators in changing landscapes for world crops. Proc. Biol. Sci. 274, 303–13.

Kremen, C., Williams, N.M., Thorp, R.W., 2002. Crop pollination from native bees at risk from agricultural intensification. Proc. Natl. Acad. Sci. U. S. A. 99, 16812–6.

Kujumgiev, A., Tsvetkova, I., Serkedjieva, Y., Bankova, V., Christov, R., Popov, S., 1999. Antibacterial, antifungal and antiviral activity of propolis of different geographic origin. J. Ethnopharmacol. 64, 235–40.

Kumazawa, S., Hayashi, K., Kajiya, K., Ishii, T., Hamasaka, T., Nakayama, T., 2002. Studies of the constituents of Uruguayan propolis. J. Agric. Food Chem. 50, 4777–82.

Kuropatnicki, A.K., Szliszka, E., Krol, W., 2013. Historical aspects of propolis research in modern times. Evid. Based. Complement. Alternat. Med. 964149.

Langenheim, J.H., 2003. Plant Resins: Chemistry, Evolution, Ecology, and Ethnobotany. Timber Press.

Langenheim, J.H., Arrhenius, S.P., Nascimento, J.C., 1981. Relationship of light intensity to leaf resin composition and yield in the tropical lenguminous gernera *Hymenaea* and *Copaifera*. Biochem. Syst. Ecol. 9(1): 27-37

Langenheim, J.H. and Hall, G.D., 1983. Sesquiterpene deterrence of a leaf-typing lepidopteran, *Stenoma ferrocanella*, on *Hymenaea stigonocarpa* in central Brazil. Biochem. Sys. Ecol. 11(1): 29-36

Leonhardt, S.D. and Bluthgen, N., 2009. A Sticky affair: Resin collection by Bornean stingless bees. Biotropica 41(6): 730-736

Leonhardt, S.D., Zeilhofer, S., Blüthgen, N., Schmitt, T., 2010. Stingless bees use terpenes as olfactory cues to find resin sources. Chem. Senses 35, 603–11.

Levin, D.A., 1976. The chemical defenses of plants to pathogens and herbivores. Annu Rev Ecol Syst 7: 121-159

Lindenfelser, L., 1967. Antimicroibal activity of propolis. Am. Bee J. 90–131.

Lindenfelser, L., 1968. *In vivo* activity of propolis against *Bacillus larvae*. J. Invertebr. Pathol. 12, 129–131.

Lokvam, J., Braddock, J., 1999. Anti-bacterial function in the sexually dimorphic pollinator rewards of *Clusia grandiflora* (Clusiaceae). Oecologia 119, 534–540.

Losey, J.E., Vaughan, M., 2006. The Economic Value of Ecological Services Provided by Insects. Bioscience 56, 311.

Mabry, T.J., Markham, K.R., Thomas, M.B., 1970. The systematic identifcation of flavonoids. Springer-Verlag, New York

Marcucci, M., 1995. Propolis: chemical composition, biological properties and therapeutic activity. Apidologie 26, 83–99.

Marcucci, M., Ferreres, F., García-Viguera, C., Bankova, V., De Castro, S., Dantas, A., Valente, P., Paulino, N., 2001. Phenolic compounds from Brazilian propolis with pharmacological activities. J. Ethnopharmacol. 74, 105–112.

Markham, K.R., 1982. Techniques of flavonoid identification. Academic Press, London

Mennerat, A., Perret, P., Bourgault, P., Blondel, J., Gimenez, O. et al., 2009. Aromatic plants in the nest of blue tits: positive effects on nestlings. Anim. Behav. 7: 569-574

Mirzoeva, O.K., Grishanin, R.N., Calder, P.C., 1997. Antimicrobial action of propolis and some of its components: the effects on growth, membrane potential and motility of bacteria. Microbiol. Res. 152, 239–46.

Mitchell, J.C., Roy, A.K., Dupuis, G., Towers, G.H.N., 1971. Allergic contact dermatitis from ragweeds (*Ambrosia* species). Arch Derm 104, 73–76.

Mohammadzadeh, S., Shariatpanahi, M., Hamedi, M., Ahmadkhaniha, R., Samadi, N., Ostad, S.N., 2007. Chemical composition, oral toxicity and antimicrobial activity of Iranian propolis. Food Chem. 103, 1097–1103.

Monzote, L., Cuesta-Rubio, O., Matheeussen, A., Van Assche, T., Maes, L., Cos, P., 2011. Antimicrobial evaluation of the polyisoprenylated benzophenones nemorosone and guttiferone A. Phytother. Res. 25, 458–62.

Morse, R., Calderone, N., 2000. The Value of Honey Bees as Pollinators of U.S. Crops in 2000. Bee Cult. 128, 1–15.

Morse, R., Flottum, K., 1997. Honey Bee Pests, Predators, and Diseases. A.I. Root Company.

Neumann, P., Elzen, P., 2004. The biology of the small hive beetle (*Aethina tumida*, Coleoptera: Nitidulidae): Gaps in our knowledge of an invasive species. Apidologie 35, 229–247.

Neumann, P., Pirk, C.W.., Hepburn, H.R., Solbrig, A.J., Ratnieks, F.L.W., Elzen, P.J., Baxter, J.R., 2001. Social encapsulation of beetle parasites by Cape honeybee colonies (*Apis mellifera capensis* Esch.). Naturwissenschaften 88, 214–216.

Nicodemo, D., Malheiros, E.B., De Jong, D., Couto, R.H.N., 2013. Increased brood viability and longer lifespan of honeybees selected for propolis production. Apidologie, DOI: 10.1007/s13592-013-0249-y

Nolkemper, S., Reichling, J., Sensch, K.H., Schnitzler, P., 2010. Mechanism of herpes simplex virus type 2 suppression by propolis extracts. Phytomedicine 17, 132–8.

NRCS, 2007. Status of pollinators in North America. National Academies Press.

Ohemeng, K., Schwender, C., Fu, K., Barrett, J., 1993. DNA gyrase inhibitory and antibacterial activity of some flavones (1). Bioorg. Med. Chem. Lett. 3, 225–230.

Oliveira, C. de, Porto, A., Bittrich, V., 1996. Floral resins of *Clusia spp.*: Chemical composition and biological function. Tetrahedron Lett. 37, 6427–6430.

Oliver, R., 2012. 2012 Almond Pollination Update. Am. Bee J. 152, 315–322.

Palo, R., 1984. Distribution of birch (*Betula spp.*), willow (*Salix spp.*), and poplar (*Populus spp.*) secondary metabolites and their potential role as chemical defense against herbivores. J. Chem. Ecol. 10.

Park, Y., Alencar, S., Aguiar, C., 2002. Botanical origin and chemical composition of Brazilian propolis. J. Agric. Food Chem. 50, 2502–2506.

Pettis, J.S., Lichtenberg, E.M., Andree, M., Stitzinger, J., Rose, R., vanengelsdorp, D., 2013. Crop pollination exposes honey bees to pesticides which alters their susceptibility to the gut pathogen Nosema ceranae. PLoS One 8, DOI: 10.1371/journal.pone.0070182.

Pinheiro, P., Justino, G., 2010. Structural Analysis of Flavonoids and Related Compounds–A Review of Spectroscopic Applications, in: Phytochemcials - A Global Prespective of Their Role in Nutrition and Health. pp. 33–56.

Plaper, A., Golob, M., Hafner, I., Oblak, M., Šolmajer, T., Jerala, R., 2003. Characterization of quercetin binding site on DNA gyrase. Biochem. Biophys. Res. Commun. 306, 530–536.

Poole, C.F., Ahmed, H., Kiridena, W., Patchett, C.C., Koziol, W.W., 2006. Revised solute descriptors for characterizing retention properties of open-tubular columns in gas chromatography and their application to a carborane-siloxane copolymer stationary phase. J. Chromatogr. A 1104, 299–312.

Popova, M.P., Chinou, I.B., Marekov, I.N., Bankova, V.S., 2009. Terpenes with antimicrobial activity from Cretan propolis. Phytochemistry 70, 1262–71.

Potts, S.G., Biesmeijer, J.C., Kremen, C., Neumann, P., Schweiger, O., Kunin, W.E., 2010. Global pollinator declines: trends, impacts and drivers. Trends Ecol. Evol. 25, 345–53.

Proksch, P., Politt, U., Wollenweber, E., Wray, V., Clark, C., 1988. Epicuticular flavonoids from *Encelia*. Planta Med. 54, 542–6.

Proksch, P., Rodriguez, E., 1983. Chromenes and benzofurans of the Asteraceae, their chemistry and biological significance. Phytochemistry 22, 2335–2348.

Reybroeck, W., Daeseleire, E., De Brabander, H.F., Herman, L., 2012. Antimicrobials in beekeeping. Vet. Microbiol. 158, 1–11.

Ricketts, T.H., Regetz, J., Steffan-Dewenter, I., Cunningham, S. a, Kremen, C., Bogdanski, A., Gemmill-Herren, B., Greenleaf, S.S., Klein, A.M., Mayfield, M.M., Morandin, L. a, Ochieng', A., Potts, S.G., Viana, B.F., 2008. Landscape effects on crop pollination services: are there general patterns? Ecol. Lett. 11, 499–515.

Rojas, A., Hernandez, L., Pereda-Miranda, R., Mata, R., 1992. Screening for antimicrobial activity of crude drug extracts and pure natural products from Mexican medicinal plants. J. Ethnopharmacol. 35, 275–283.

Rousi, M., Tahvanainen, J., Uotila, I., 1991. A mechanism of resistance to hare browsing in winter-dormant European white birch (*Betula pendula*). Am. Nat. 137, 64–82.

Seeley, T., Morse, R., 1976. The nest of the honey bee (*Apis mellifera L.*). Insectes Soc. 23, 495–512.

Schmid-Hempel, P., 1998. Parasites in social insects. Princeton University Press.

Sforcin, J.M., 2007. Propolis and the immune system: a review. J. Ethnopharmacol. 113, 1–14.

Sforcin, J.M., Bankova, V., 2011. Propolis: is there a potential for the development of new drugs? J. Ethnopharmacol. 133, 253–60.

Shuler, R., Roulston, T., Farris, G., 2005. Farming practices influence wild pollinator populations on squash and pumpkin. J. Econ. Entomol. 98, 790–795.

Simone, M., Evans, J.D., Spivak, M., 2009. Resin collection and social immunity in honey bees. Evolution 63, 3016–22.

Simone-Finstrom, M., Gardner, J., Spivak, M., 2010. Tactile learning in resin foraging honeybees. Behav. Ecol. Sociobiol. 64(10): 1609-1617

Simone-Finstrom, M., Spivak, M., 2010. Propolis and bee health: the natural history and significance of resin use by honey bees. Apidologie 41, 295–311.

Simone-Finstrom, M.D., Spivak, M., 2012. Increased resin collection after parasite challenge: a case of self-medication in honey bees? PLoS One 7, DOI: 10.1371/journal.pone.0034601

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., Siuzdak, G., 2006. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal. Chem. 78: 779-787

Spivak, M., Gilliam, M., 1998. Hygienic behaviour of honey bees and its application for control of brood diseases and varroa. Bee World 79, 169–186.

Spivak, M., Reuter, G., 2001. Resistance to American foulbrood disease by honey bee colonies *Apis mellifera* bred for hygienic behavior. Apidologie 32, 555–565.

Srivastava, R., Proksch, P., Wray, V., 1990. Toxicity and antifeedant activity of a sesquiterpene lactone from *Encelia* against *Spodoptera littoralis*. Phytochemistry 29, 3445–3448.

Stapleton, P.D., Shah, S., Hamilton-Miller, J.M.T., Hara, Y., Nagaoka, Y., Kumagai, A., Uesato, S., Taylor, P.W., 2004. Anti-*Staphylococcus aureus* activity and oxacillin resistance modulating capacity of 3-O-acyl-catechins. Int. J. Antimicrob. Agents 24, 374–80.

Sumner, D., Boriss, H., 2006. Bee-conomics and the leap in pollination fees. Agric. Resour. Econ. Updat. 9, 9–11.

Sun, Y., Zhang, R., Wang, Q., Xu, B., 1993. Programmed-temperature gas chromatographic retention index. J. Chromatogr. A 657, 1–15.

Taipale, H., Vepsäläinen, J., Laatikainen, R., Reichardt, P., Lapinjoki, S., 1993. Isolation and structure determination of three triterpenes from bark resin of juvenile european white birch. Phytochemistry 34, 755–758.

Tautenhahn, R., Böttcher, C., Neumann, S., 2008. Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics 9(504), doi: 10.1186/1471-2105-9-504

Teixeira, E.W., Negri, G., Meira, R.M.S.A., Message, D., Salatino, A., 2005. Plant origin of green propolis: Bee behavior, plant anatomy and chemistry. eCAM 2(1): 85-92

Tomás-Barberán, F., García-Viguera, C., 1993. Phytochemical evidence for the botanical origin of tropical propolis from Venezuela. Phytochemistry 34, 191–196.

Trygg, J., Holmes, E., Lundstedt, T., 2007. Chemometrics in Metabolomics. J. Proteome Res. 6(2), p. 469-79

USDA, NRCS National Plant Data Team (2013) USDA PLANTS Database, plants.usda.gov. Accessed Sept 16, 2013.

USDA NASS, 2013a. 2013 California Almond Objective Measurement Report.

USDA NASS, 2013b. 2012 California Almond Acreage Report.

Valkama, E., Salminen, J., Koricheva, J., Pihlaja, K., 2003. Comparative Analysis of Leaf Trichome Structure and Composition of Epicuticular Flavonoids in Finnish Birch Species. Ann. Bot. 91, 643–655.

Valkama, E., Salminen, J.-P., Koricheva, J., Pihlaja, K., 2004. Changes in leaf trichomes and epicuticular flavonoids during leaf development in three birch taxa. Ann. Bot. 94, 233–42.

van den Dool, H., Kratz, P.D., 1963. A generalization of the retention index system including linear temperature programmed gas—liquid partition chromatography. J. Chromatogr. A 463–471.

vanEngelsdorp, D., Caron, D., Hayes, J., Underwood, R., Henson, M., Rennich, K., Spleen, A., Andree, M., Snyder, R., Lee, K., Roccasecca, K., Wilson, M., Wilkes, J., Lengerich, E., Pettis, J., 2012. A national survey of managed honey bee 2010-11 winter colony losses in the USA: results from the Bee Informed Partnership. J. Apic. Res. 51, 115–124.

vanengelsdorp, D., Meixner, M.D., 2010. A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. J. Invertebr. Pathol. 103 Suppl , S80–95.

vanengelsdorp, D., Evans, J.D., Saegerman, C., Mullin, C., Haubruge, E et al., 2009. Colony collapse disorder: A descriptive study. PLoS One 4, doi: 10.1371/journal.pone.0006481

Vezzani, S., Moretti, P., Castello, G., 1997. Automatic prediction of retention times in multi-linear programmed temperature analyses. J. Chromatogr. A 767, 115–125.

Wallace, H.M., Lee, D.J., 2010. Resin-foraging by colonies of *Trigona sapiens* and *T. hockingsi* (Hymenoptera : Apidae , Meliponini) and consequent seed dispersal of *Corymbia torelliana* (Myrtaceae)*. Apidologie 41, 428–435.

Wallace, H.M., Trueman, S.J., 1995. Dispersal of *Eucalyptus torelliana* seeds by the resin-collecting stingless bee, *Trigona carbonaria*. Oecologia 104, 12–16.

Wilson, M.B., Spivak, M., Hegeman, A.D., Rendahl, A., Cohen, J.D., 2013. Metabolomics Reveals the Origins of Antimicrobial Plant Resins Collected by Honey Bees. PLoS One 8, DOI: 10.1371/journal.pone.0077512

Wilson-Rich, N., Spivak, M., Fefferman, N.H., Starks, P.T., 2009. Genetic, individual, and group facilitation of disease resistance in insect societies. Annu. Rev. Entomol. 54: 405-423

Witham, T.G., 1983. Host manipulation of parasites: within plant variation as a defense against rapidly evolving pests. Variable Plants and Herbivores in Natural and Managed Systems, New York: Academic Press. p. 15-41

Wollenweber, E., Buchmann, S.L., 1997. Feral honey bees in the Sonoran Desert: Propolis sources oter than Poplars (*Populus spp.*). Z. Naturforsch. C. 52c, 530–535.

Wollenweber, E., Dietz, V., 1981. Occurrence and distribution of free flavonoid aglycones in plants. Phytochemistry 20, 869–932.

Wollenweber, E., Hradetzky, D., Mann, K., Roitman, J.N., Yatskievych, G., Proksch, M., Proksch, P., 1987. Exudate Flavonoids from Aerial Parts of Five *Ambrosia* Species. J. Plant Physiol. 131, 37–43.

Yamamoto, S., Otto, A., Simoneit, B.R.T., 2004. Lignans in resin of Araucaria angustifolia by gas chromatography/mass spectrometry. J. Mass Spectrom. 39, 1337–47.

Zellner, B., Bicchi, C., Dugo, P., 2008. Linear retention indices in gas chromatographic analysis: a review. Flavour and Fragrance Journal 23, 297–314.

Zhao, C.-X., Zhang, T., Liang, Y.-Z., Yuan, D.-L., Zeng, Y.-X., Xu, Q.-S., 2007. Conversion of programmed-temperature retention indices from one set of conditions to another. J. Chromatogr. A 1144, 245–54.

Appendix A

Experimental factors affecting the accuracy of retention time projection in GC-MS

In collaboration with Dr. Paul Boswell, University of Minnesota

Summary

The systematic annotation of known metabolites is a major bottle-neck in mass spectrometry-based metabolomics and natural products chemistry. Retention data could be a powerful tool in identifying "known unknown" compounds; however, current methods for documenting retention information are too dependent on the particular instrument or experiential conditions used, making retention databases only marginally useful to the community. "Retention projection" is a new methodology that overcomes the limitations of traditional linear retention indexing in GC-MS by accounting for both experimental and unintentional differences between user-generated data collected with different methods or by labs and instruments by very precisely modeling instrument behavior and subsequently predicting compound retention times from isothermal data. We have previously shown that retention projection is robust under changing temperature programs, flow rates, and inlet pressures and here we continue to assess its ability to accurately predict retention times with changing column lengths, phase ratios, and 16 other injection, inlet, and MS variables. We found that using a very short column (<15 m), having a dirty liner/column, and column overloading detrimentally affected retention projection accuracy. We have used this data to create troubleshooting guidelines for the open source retention projection software at www.retentionprediction.org/gc.

Introduction

Capturing retention data is a powerful way to leverage chemical information for making metabolite annotations in biological samples by GC-MS. Utilizing both mass and retention information together greatly increases one's ability to identify known compounds relative to using accurate mass or fragmentation pattern alone, even when

retention information is used in combination with unit mass resolution (Boswell et al., 2011). Experimentally, the community wastes much time and effort to identify known compounds in biological samples when authentic standards are not available. In Chapter 4, it took LC-MS-MS, NMR, and UV-Vis analysis to unambiguously identify 5-methoxyl pinobanksin (a known compound) because an authentic standard was not available.

Unfortunately, current methodologies (e.g. linear retention indexing) only capture retention information in such a way that data is difficult or impossible to reproduce across different methods, instruments, or laboratories, especially if temperature programmed GC is used (Sun et al., 1993; Zellner et al., 2008; Zhao et al., 2007). Linear retention indexing assumes fixed temperature vs. retention (log k') relationships regardless of experimental conditions, which is almost never true (Barnes et al., 2013; Boswell et al., 2012) and so RI values are only consistent as long as the same instrument and temperature-program are used. Overall, this variability under differing experimental conditions makes libraries of RI values unreliable community-based tools for metabolite identification (Sun et al., 1993; Zellner et al., 2008; Zhao et al., 2007). Even if universal methods were adopted, unintentional differences between instruments, such as temperature calibration, are sufficient to cause significant changes in RI values (Barnes et al., 2013).

Furthermore, shared retention index databases offer no way to determine the appropriate retention time tolerance window to use with each retention index. For the purpose of compound identification, retention time tolerance windows are arguably just as important as predicted retention times – they make it possible to exclude a possible identity for a peak based on its retention time *with a known level of confidence*. Since retention indices cannot account for unintentional differences between GC systems, the appropriate retention time tolerance windows vary from lab to lab. Therefore, one must have a standard of each analyte physically on hand to confirm its identity, which is impossible for the vast majority of biological compounds. Overall, retention indexing is experimentally restrictive, imprecise, and offers no clear level of confidence to interpret the information it provides.

"Retention projection" is a new retention data-capturing methodology that overcomes the limitations of retention indexing and is freely available at checkyourgc.org and retentionprediction.org. It overcomes these limitations by accounting for both experimental and unintentional differences between user-generated data collected with different methods or by labs and instruments through very precise modeling of instrument behavior and subsequently predicting compound retention times from isothermal data. This modeling of instrument behavior is achieved by using a series of n-alkane standards with known temperature vs. log k' relationships to back-calculate the actual instrument behavior *that must have occurred* to produce the experimental n-alkane retention times (Boswell et al., 2012). Simultaneously, this software uses the modeled temperature vs. time and hold-up time vs. temperature profiles to very accurately predict the retention times of other compounds from a database of isothermal data (Boswell et al., 2012; Vezzani et al., 1997). Predicted retention times are most often within 1 second of experimental retention times across changing temperature programs, flow rates, inlet pressures, instruments, and laboratories. Since accuracy is lab-independent (Barnes et al., 2013), it is also possible to calculate tolerance windows for each predicted retention time to assess if that compound truly appears in the sample without having an authentic standard physically on hand (Barnes et al., 2013).

The ultimate utility of retention projection depends on this methodology accurately predicting retention times under changing experimental conditions. We have previously shown that retention projection is robust against changing instruments, temperature programs, flow rates, and inlet pressures (Barnes et al., 2013; Boswell et al., 2012), and here we assess retention projection accuracy under changing column lengths, phase ratios, and 16 other injection, inlet, and MS variables. These experiments represent a test of the practical robustness of retention projection methodology to a range of practical experimental conditions not specifically tested in the multi-lab study.

Materials and Methods

Measuring retention prediction accuracy

Retention projection accuracy was measured as previously described (Barnes et al., 2013; Boswell et al., 2012). Briefly, 12 test compounds were selected as models for the five most prominent molecular interactions leading to retention in GC, as represented by Abraham descriptors (Abraham et al., 2004). These standards included hydrogen bond donors (phenol, resorcinol, 1-naphthol), hydrogen bond acceptors (*N*,*N*-dimethylisobutyramide, benzamide, dextromethorphan), compounds that interact by π/lone pair electrons (ethylbenzene, naphthalene, anthracene), and compounds that interact by dipole-dipole or dipole-induced dipole interactions (*N*,*N*-diethylacetamide, 4-nitroaniline, caffeine) (Atapattu and Poole, 2008; Poole et al., 2006). All compounds vary in their gas-liquid partition coefficients (Atapattu and Poole, 2008), and elute over a wide range of retention times (Barnes et al., 2013). The 12 test compounds plus 25 n-alkanes for the back-calculation of instrument behavior were dissolved in ethyl acetate at 100 μM concentration. Chemicals and solvents were purchased from Sigma-Aldrich (St. Louis, MO), Alfa Aesar (Ward Hill, MA), or TCI America (Portland, OR).

Overall accuracy of retention projections were taken as the root mean squared difference between experimental and projected retention times for the set of 12 test compounds (Barnes et al., 2013).

Calculation of retention indices

Linear retention indices were calculated for comparison with retention projection following standard methods(Van den Dool and Kratz, 1963). To show changes in the accuracy of retention indices with changing experimental conditions, differences between indices calculated for the first, typically standard, condition (e.g. 30 m column) and the treatment condition (e.g. 25 m column) were compared to determine retention time prediction accuracy.

Instrumentation

Experiments were performed with a HP 5970 GC coupled to a HP 5890 single quadrupole MS or Thermo GC Trace Ultra coupled to a Thermo TSQ Quantum triple quadrupole MS. The carrier gas was UPH helium (Matheson) scrubbed for oxygen, water, and hydrocarbons with a R&D Separations QC+ system at a 1:10 split ratio. Restek straight, split/splitless, IP deactivated quartz liners 2 x 6.5 x 78.5 mm were used. Agilent DB-5MS UI columns with 0.25 mm i.d. and 0.25 µm film thickness were used as the standard condition. Column i.d. and film thickness were varied to 0.32 mm and 1 µm, respectively, as noted.

Column exposure to oxygen

The GC-MS system was connected to UHP helium spiked with 6.2 ppm oxygen (Oxygen Services) for oxygen exposure experiments. The system's oxygen scrubber was removed, though water and hydrocarbon scrubbers remained in-line for experiments.

Phase ratio studies

Phase ratios for the 0.32 mm i.d. column and 1 µm film thickness column were tested by taking retention data and manipulating column i.d. and film thickness input parameters in the retention projection software. Parameters producing the most accurate retention projections were regarded as the actual column inner diameter and film thickness.

Preparation of urine samples

20 mL urine was adjusted to basic pH conditions by adding approximately 2 mL of 1M NaOH. 10 mL of methylene chloride was added, and the sample was centrifuged at 750 x *g* for 5 minutes. The aqueous layer (top) was discarded. The organic layer (bottom) was transferred to a new vial, evaporated to dryness in a 65 ^{o}C water bath, and reconstituted with 1 mL of ethyl acetate.

Results and Discussion

Malfunctioning GC oven

Aluminum blocks were placed in the GC oven in order to simulate a malfunction causing changes in the rate of GC oven heating. Deviations from the "ideal" or programmed rate of heating are clearly evident at higher temperatures (Fig 1). The temperature profile actually produced by the GC was highly distorted, but the accuracy of retention projections was unchanged with an accuracy of ±0.43 s without blocks in the oven and ±0.35 s with blocks in the oven.

Comparison of retention indexing and retention projection with changing column length and phase ratio.

Previously, we showed that retention projection is up to 160-fold more accurate than linear retention indexing when varying temperature program, flow rate, or inlet pressure across different labs choosing their own unique experimental methods (Barnes et al., 2013). Table 1 compares the accuracies of retention projection and retention indexing when column length, i.d., and film thickness are varied. For differing column lengths, retention projection was much more accurate than linear indexing with a deviation of ±0.66 sec vs. a deviation of ±9.4 sec for linear indexing for the 15 m column (column length = 30 m in the standard condition): however, we did see a slight decrease in retention projection accuracy with decreasing column length. Retention projection accuracy was also much more accurate than linear indexing for changing column film thickness and i.d., respectively (Table 1). Table 1 shows that retention project will be robust in applications that use columns shorter than 30 m and different phase ratios. This was important because all isothermal data used to predict retention times in the database were collected using a 30 m column.

Surprisingly, we had to alter the phase ratio input parameters (column i.d. and film thickness) in our retention projection software from those stated by the manufacturer to get the most accurate retention predictions when using the 0.32 i.d. and 1 µm film thickness columns (Fig 2). Since accuracy for retention projection in this case is purely a

function of these columns having the manufacture's stated phase ratio, the data in Fig 2 lead us to believe that phase ratios not stated accurately enough to make the most accurate retention projections (Agilent Technologies, personal communication); however we are confident that phase ratios are consistent between columns. Our software has been updated to use the optimized phase ratio values by automatically acknowledging Agilent DB-5MS UI 0.32 mm inner diameter columns and 1 µm film thickness columns as 0.36 mm i.d. columns and 0.9 µm film thickness columns, respectively.

Accuracy of retention projection with changing practical experimental variables

Retention projection accuracy was robust against changes in a large number of practical experimental variables. Injection variables (manual injection speed, split/splitless injection) did not affect retention projection accuracy, so long as concentration was held constant (Table 2). The cold trapping phase of a good gradient method (5 mins at 60 ºC in our experiments) likely minimizes any small effects caused by different rates of solvent/analyte vaporization that correspond to our different injection methods. Inlet variables also had no effect on accuracy when inlet temperature, column cut, column length in the inlet, or liners and liner packing materials were varied (Table 2). A leaky septum or using poorly purified carrier gases could introduce oxygen into a hot GC-MS system and cause column oxidation over time, but prolonged exposure to low-levels of oxygen did not have any effect on retention projection accuracy (Table 2). Lastly, variations in MS source temperature had no effect on accuracy (Table 2), and all compounds remained detectable at all temperatures. Any one of the above mentioned factors may be accidently in the wrong configuration during a real analysis, but users can still have confidence that retention projection will perform properly and accurately under accidental changes.

Three variables among those tested were found to decrease retention projection accuracy (Table 3). Accuracy decreased with increasing sample concentration (Table 3) suggesting that best results are achieved by making sure that only 0.1 nmol of each test mix compound are put on the column during a given method for column suitability

checks performed with checkyourgc.org, and that overloading the column with other analytes would give non-optimal results. This could be particularly evident when switching from a split to a splitless method, where it is easy to forget about the dilution factor applied by split injections. Secondly, setting the transfer line temperature far below the maximum of our temperature program (160 ºC vs. 320ºC in our experiments) had a slight negative effect on accuracy (Table 3). Lastly, repeated injections of dirty samples have a relatively large negative effect on accuracy, likely due to spurious retention in the liner or changes in column selectivity. Dirty injections were by far the most detrimental of the three negative factors (Table 3). If a column is damaged to the point where cutting is required, we demonstrated above that retention projection will be just as accurate with the shorter column (Table 1). We have used this data to create troubleshooting guidelines for the open source retention projection software at www.retentionprediction.org/gc and updated checkyourgc.org to alert users that if their system fails the suitability check, the above mentioned factors are likely the cause.

Conclusions

The retention projection methodology proved to be a robust way to share gas chromatographic retention data and confirmed its suitability for application as a community tool for metabolite annotation in GC-MS. Previous work showed that the accuracy is unaffected (relative to that expected from theory) by the temperature program, the flow rate, the inlet pressure, and the GC instrument used. Here, we tested 18 other experimental factors and found that only 5 could be responsible for reduced accuracy in retention projections: an overloaded column, a low transfer line temperature, prior injection of dirty samples, and use of a very short column ($\leq$ 15 m).

In addition, the methodology was shown to account for severe temperature program errors, differences in film thickness, differences in inner diameter, and to some extent, differences in column length. Retention projections were 10- to 30-fold more accurate than retention indices under the latter three conditions. For changing film thickness and inner diameter, we had to account for small, but likely consistent

differences in the actual film thickness and inner diameters of the DB-5MS UI column from stated values. The online retention projection software at www.retentionprediction.org/gc was modified to automatically take the differences into account. Our future goal is to building a much larger database of isothermal retention data that can be used as a community tool for GC-MS metabolite annotation in biological samples.

Figure 1 – (A) Photograph of the GC oven containing two Al blocks (circled). (B) Comparison of back-calculated profiles with and without the blocks in the oven. When the blocks were placed into the oven, the temperature profile actually produced by the GC was highly distorted, but the accuracy of retention projections was unaffected.

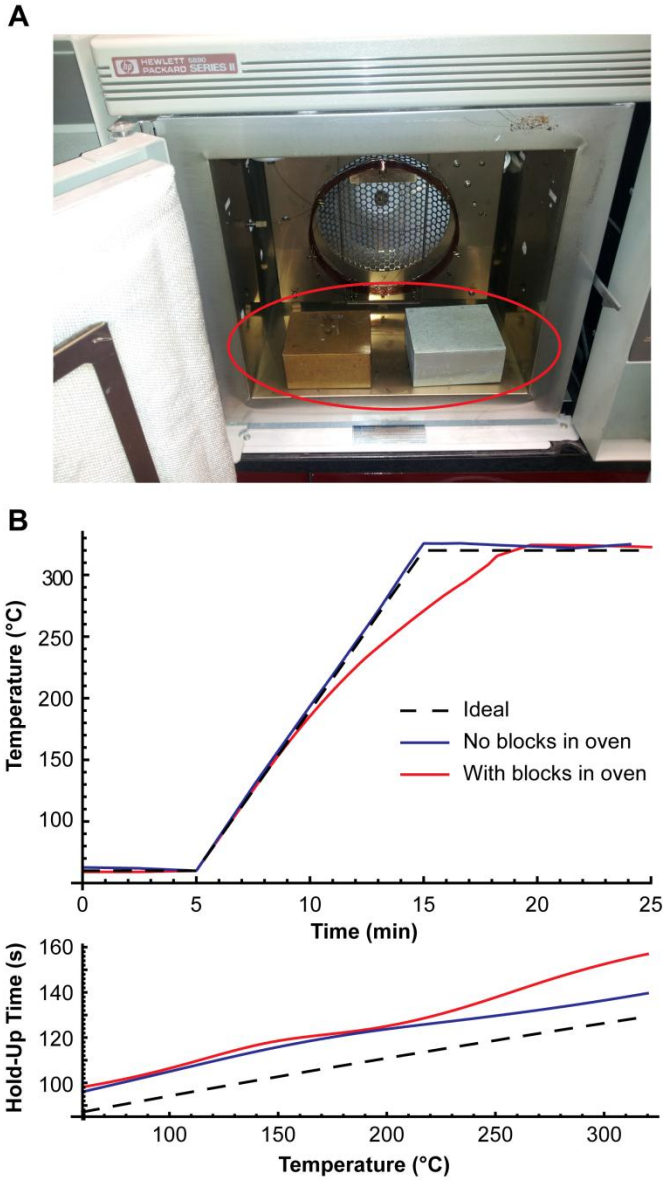


Figure 2 (A, B) – Optimization of retention projection accuracy in 0.32 mm i.d. and 1 µm film thickness Agilent DB-5MS UI columns. Retention projection accuracy was optimized for columns with non-standard phase ratios by changing the input column i.d. for the 0.32 mm i.d. column (A) or input film thickness for the 1 µm film thickness column (B). The low point in the parabolic graph represents the optimum or "true" column i.d. or film thickness.

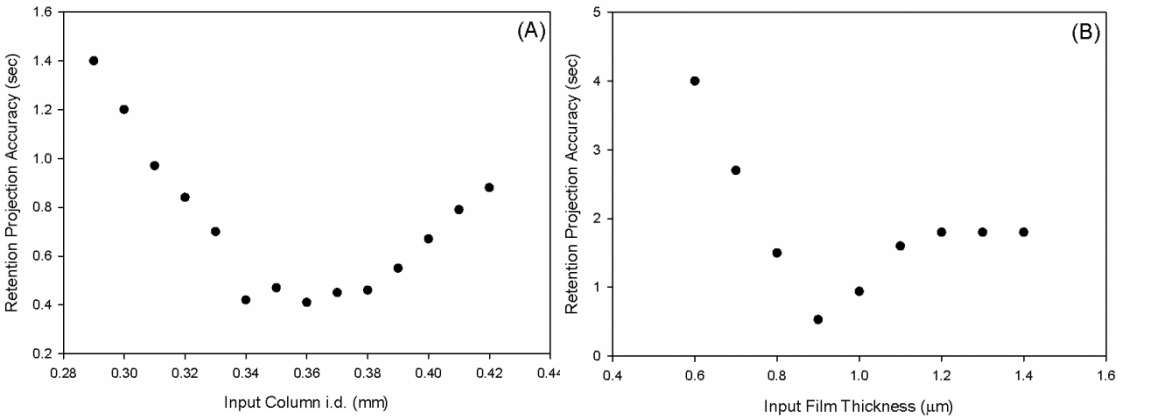


Table 1 – Effect of column length, film thickness, and column i.d. on the accuracy of retention projection and linear indexing. Retention projection accuracy was determined as the root mean squared difference between experimental and projected retention times for the set of 12 test compounds. Linear indexing accuracy is always relative to the standard condition (e.g. linear indexing accuracy for the 25 m column is 4.9 sec less accurate when it assumes a column length of 30 m), because this is an inherent property of predicting retention times with linear indexing. * indicate that these values are optimized (see Fig 2).

Column Length	**Accuracy of RT Projection**	**Accuracy of Linear Indexing**
30 m	±0.44 sec	---
25 m	±0.45 sec	±4.9 sec
20 m	±0.55 sec	±5.8 sec
15 m	±0.66 sec	±9.4 sec
Film Thickness		
0.25 µm	±0.44 sec	---
1 µm	±0.53 sec*	±15.7 sec
Column i.d.		
0.25 mm	±0.44 sec	---
0.32 mm	±0.41 sec*	±14.5 sec

Table 2 – Experimental variables not affecting retention projection accuracy. Accuracy was determined as the root mean squared difference between experimental and projected retention times for the set of 12 test compounds. Detailed descriptions of each experiment can be found in the methods section.

Manual Injection Speed	**Accuracy of RT projection**
Fast (< 0.5 sec)	±0.42 sec
Slow (5 sec)	±0.44 sec
Split or Splitless Injection	
Split (1:10)	±0.42 sec
Splitless	±0.43 sec
Inlet Temperature	
300ºC	±0.44 sec
290 ºC	±0.42 sec
280 ºC	±0.44 sec
270 ºC	±0.43 sec
250 ºC	±0.48 sec
230 ºC	±0.45 sec
200 ºC	±0.44 sec
Column cut	
Straight (good)	±0.42 sec
Angled (bad)	±0.45 sec
Length of Column in Inlet	
2 mm	±0.44 sec
1.2 mm	±0.39 sec
2.2 mm	±0.37 sec
Inlet Liner	
Agilent, 4 mm gooseneck, glass	±0.44 sec
SGE, 4 mm straight, glass	±0.38 sec
Agilent, 2 mm straight splitless, quartz	±0.49 sec
Restek Siltek, 2 mm straight splitless, quartz	±0.45 sec
Restek Sky, 4 mm straight, glass, packed w/ deactivated glass wool	±0.40 sec
Supelco, 4 mm straight, glass, packed w/ deactived glass wool	±0.40 sec

Column Conditioning	
None	±0.42 sec
3 hrs @ 320ºC	±0.45 sec
Oxygen Exposure	
Before	±0.42 sec
After (3 days 6.2 ppm O_2 @ 300ºC)	±0.39 sec
MS Source Temperature	
300 ºC	±0.52 sec
220 ºC	±0.50 sec
180 ºC	±0.49 sec
100 ºC	±0.52 sec

Table 3 – Experimental variables that negatively affected retention projection accuracy. Accuracy was determined as the root mean squared difference between experimental and projected retention times for the set of 12 test compounds. Detailed descriptions of each experiment can be found in the methods section.

Test Mix Concentration	**Accuracy of RT projection**
0.1 nmol	±0.42 sec
1 nmol	±0.68 sec
Sample Exposure	
Before samples	±0.51 sec
After 7 dirty urine injections	±1.28 sec
Transfer Line Temperature	
320ºC	±0.44 sec
240 ºC	±0.57 sec
160 ºC	±0.66 sec

Appendix B

Compounds reported in North American poplar resins, *Populus spp.*

This appendix is compiled from data appearing in the following publications:

(1) English, S., Greenaway, W., Whatley, F.R. (1991). Analysis of *Populus trichocarpa* bud exudate by GC-MS. Phytochemistry 30(2), 531-533

(2) English, S., Greenaway, W., Whatley, F.R. (1992). Analysis of phenolics in the bud exudates of *Populus deltoides, P. fremontii, P. sargentii* and *P. wislizenii* by GC-MS. Phytochemistry 31(4), 1255-1260

(3) Greenaway, W and Whatley, F.R. (1990). Analysis of phenolics of bud exudate of *Populus angustifolia* by GC-MS. Phytochemistry 29(8), 2551-2554

(4) Isidorov, VA and Vinogorova, VT (2003). GC-MS analysis of compounds extracted from buds of *Populus balsamifera* and *Populus nigra*. Z. Naturforsch, 58c, 355-360

*Compound identities were determined by GC-MS fragmentation analysis

Publication	Species	Location	Compound Name	Exact mass
4	**P. balsamifera**	**Białystok, Poland**	**2-propen-1-one**	**56.026215**
4	**P. nigra**	**Białystok, Poland**	**2-propen-1-one**	**56.026215**
1	**P. trichocarpa**	**Alaska, USA**	**benzyl alcohol**	**108.05752**

4	**P. balsamifera**	**Białystok, Poland**	**benzyl alcohol**	**108.05752**
1	**P. trichocarpa**	**Alaska, USA**	**hydroquinone**	**110.03678**
3	**P. angustifolia**	**S. Alberta, Canada**	**1,2-benzenediol**	**110.03678**
4	**P. balsamifera**	**Białystok, Poland**	**heptanal**	**114.10447**
1	**P. trichocarpa**	**Alaska, USA**	**benzoic acid**	**122.03678**
1	**P. fremontii**	**Arizona, USA**	**benzoic acid**	**122.03678**
1	**P. trichocarpa**	**Alaska, USA**	**4-hydroxybenzaldehyde**	**122.03678**
3	**P. angustifolia**	**S. Alberta, Canada**	**benzoic acid**	**122.03678**
4	**P. balsamifera**	**Białystok, Poland**	**salicylaldehyde**	**122.03678**
1	**P. trichocarpa**	**Alaska, USA**	**2-phenylethanol**	**122.07317**
2	**P. trichocarpa**	**Alaska, USA**	**1-phenol ethanol**	**122.07317**
3	**P. angustifolia**	**S. Alberta, Canada**	**2-phenylethanol**	**122.07317**
4	**P. balsamifera**	**Białystok, Poland**	**2-phenylethanol**	**122.07317**
4	**P. nigra**	**Białystok, Poland**	**2-phenylethanol**	**122.07317**
3	**P. angustifolia**	**S. Alberta, Canada**	**1,2,3-trihydroxybenzene**	**126.0317**
1	**P. trichocarpa**	**Alaska, USA**	**cinnamyl alcohol**	**134.07317**

1	**P. trichocarpa**	**Alaska, USA**	**4-acetophenol**	**136.05243**
1	**P. trichocarpa**	**Alaska, USA**	**4-hydroxybenzoic acid**	**138.0317**
3	**P. angustifolia**	**S. Alberta, Canada**	**3,4-dihydroxybenzaldehyde**	**138.0317**
1	**P. trichocarpa**	**Alaska, USA**	**cinnamic acid**	**148.05243**
3	**P. angustifolia**	**S. Alberta, Canada**	**cinnamic acid**	**148.05243**
1	**P. trichocarpa**	**Alaska, USA**	**hydrocinnamic acid**	**150.06808**
1	**P. trichocarpa**	**Alaska, USA**	**4-coumaryl alcohol**	**150.06808**
3	**P. angustifolia**	**S. Alberta, Canada**	**hydrocinnamic acid**	**150.06808**
4	**P. nigra**	**Białystok, Poland**	**ethyl benzoate**	**150.06808**
1	**P. trichocarpa**	**Alaska, USA**	**4-anisic acid**	**152.04735**
1	**P. trichocarpa**	**Alaska, USA**	**3,4-dihydroxyacetophenone**	**152.04735**
3	**P. angustifolia**	**S. Alberta, Canada**	**3,4-dihydroxybenzoic acid**	**154.02661**
4	**P. nigra**	**Białystok, Poland**	**1,8-cineole**	**154.13577**
1	**P. trichocarpa**	**Alaska, USA**	**4-coumaric acid**	**164.04735**
3	**P. angustifolia**	**S. Alberta, Canada**	**4-coumaric acid**	**164.04735**
4	**P. balsamifera**	**Białystok, Poland**	**eugenol**	**164.08373**

4	**P. nigra**	**Białystok, Poland**	**eugenol**	**164.08373**
3	**P. angustifolia**	**S. Alberta, Canada**	**5-phenylpenta-2:4-dienoic acid**	**174.06808**
4	**P. balsamifera**	**Białystok, Poland**	**(E)-ethyl cinnamate**	**176.08373**
3	**P. angustifolia**	**S. Alberta, Canada**	**phenylpentenoic acid**	**178.09938**
2	**P. deltoides**	**Wisconsin, USA**	**caffeic acid**	**180.04226**
3	**P. angustifolia**	**S. Alberta, Canada**	**caffeic acid**	**180.04226**
1	**P. trichocarpa**	**Alaska, USA**	**4-methoxyhydrocinnamic acid**	**180.07865**
2	**P. deltoides**	**Wisconsin, USA**	**isoferulic acid**	**194.05791**
2	**P. deltoides**	**Wisconsin, USA**	**ferulic acid**	**194.05791**
3	**P. angustifolia**	**S. Alberta, Canada**	**isoferulic acid**	**194.05791**
3	**P. angustifolia**	**S. Alberta, Canada**	**ferulic acid**	**194.05791**
4	**P. balsamifera**	**Białystok, Poland**	**γ-curcumene**	**202.17215**
4	**P. balsamifera**	**Białystok, Poland**	**Ar-curcumene**	**202.17215**
4	**P. balsamifera**	**Białystok, Poland**	**(Z)-β-farnesene**	**204.1878**
4	**P. balsamifera**	**Białystok, Poland**	**(E)-β-farnesene**	**204.1878**
4	**P. balsamifera**	**Białystok, Poland**	**β-Selinene**	**204.1878**

4	**P. balsamifera**	**Białystok, Poland**	**δ-guaiene (α-bulnesene)**	**204.1878**
4	**P. balsamifera**	**Białystok, Poland**	**β-bisabolene**	**204.1878**
4	**P. balsamifera**	**Białystok, Poland**	**(Z)-caryophyllene**	**206.20345**
4	**P. nigra**	**Białystok, Poland**	**(Z)-caryophyllene**	**206.20345**
4	**P. balsamifera**	**Białystok, Poland**	**(E)-caryophyllene**	**206.20345**
3	**P. angustifolia**	**S. Alberta, Canada**	**2-methyl-2-propenyl-(E)-caffeate**	**208.07356**
1	**P. trichocarpa**	**Alaska, USA**	**benzyl benzoate**	**212.08373**
4	**P. balsamifera**	**Białystok, Poland**	**caryophyllene oxide**	**220.18272**
4	**P. nigra**	**Białystok, Poland**	**caryophyllene oxide**	**220.18272**
4	**P. balsamifera**	**Białystok, Poland**	**Guaiol**	**222.19837**
4	**P. nigra**	**Białystok, Poland**	**Guaiol**	**222.19837**
4	**P. balsamifera**	**Białystok, Poland**	**γ-eudesmol, 10-epi-**	**222.19837**
4	**P. nigra**	**Białystok, Poland**	**γ-eudesmol, 10-epi-**	**222.19837**
4	**P. balsamifera**	**Białystok, Poland**	**β-budesmol (β-selinenol)**	**222.19837**
4	**P. nigra**	**Białystok, Poland**	**β-budesmol (β-selinenol)**	**222.19837**
4	**P. balsamifera**	**Białystok, Poland**	**α-eudesmol**	**222.19837**

4	**P. nigra**	**Białystok, Poland**	**α-eudesmol**	**222.19837**
4	**P. balsamifera**	**Białystok, Poland**	**bulnesol**	**222.19837**
4	**P. nigra**	**Białystok, Poland**	**bulnesol**	**222.19837**
1	**P. trichocarpa**	**Alaska, USA**	**1-phenylethyl benzoate**	**226.09938**
1	**P. trichocarpa**	**Alaska, USA**	**benzyl salicylate**	**228.07865**
1	**P. trichocarpa**	**Alaska, USA**	**salicyl benzoate**	**228.07865**
1	**P. trichocarpa**	**Alaska, USA**	**4-hydroxybenzyl benzoate**	**228.07865**
1	**P. trichocarpa**	**Alaska, USA**	**(E)-cinnamyl benzoate**	**238.09938**
1	**P. trichocarpa**	**Alaska, USA**	**benzyl-2,5-dihydroxybenzoate**	**244.07356**
2	**P. deltoides**	**Wisconsin, USA**	**3-methyl-3-butenyl-(E)-caffeate**	**248.10486**
2	**P. deltoides**	**Wisconsin, USA**	**2-methyl-2-butenyl-(E)-caffeate**	**248.10486**
2	**P. deltoides**	**Wisconsin, USA**	**3-methyl-2-butenyl-(E)-caffeate**	**248.10486**
3	**P. angustifolia**	**S. Alberta, Canada**	**3-methyl-3-butenyl-(E)-caffeate**	**248.10486**
3	**P. angustifolia**	**S. Alberta, Canada**	**2-methyl-2-butenyl-(E)-caffeate**	**248.10486**
3	**P. angustifolia**	**S. Alberta, Canada**	**3-methyl-2-butenyl-(E)-caffeate**	**248.10486**
3	**P. angustifolia**	**S. Alberta, Canada**	**2-methyl-2-propenyl-(E)-isoferulate**	**248.10486**

1	**P. fremontii**	**Arizona, USA**	**chrysin**	**254.05791**
2	**P. deltoides**	**Wisconsin, USA**	**chrysin**	**254.05791**
3	**P. angustifolia**	**S. Alberta, Canada**	**chrysin**	**254.05791**
1	**P. trichocarpa**	**Alaska, USA**	**(E)-cinnamyl-4-hydroxybenzoate**	**254.0943**
1	**P. trichocarpa**	**Alaska, USA**	**benzyl-(E)-4-coumarate**	**254.0943**
3	**P. angustifolia**	**S. Alberta, Canada**	**benzyl-(E)-4-coumarate**	**254.0943**
1	**P. trichocarpa**	**Alaska, USA**	**pinocembrin**	**256.07356**
1	**P. fremontii**	**Arizona, USA**	**pinocembrin**	**256.07356**
1	**P. trichocarpa**	**Alaska, USA**	**pinocembrin chalcone**	**256.07356**
1	**P. fremontii**	**Arizona, USA**	**pinocembrin chalcone**	**256.07356**
2	**P. deltoides**	**Wisconsin, USA**	**pinocembrin**	**256.07356**
2	**P. deltoides**	**Wisconsin, USA**	**pinocembrin chalcone**	**256.07356**
3	**P. angustifolia**	**S. Alberta, Canada**	**pinocembrin**	**256.07356**
3	**P. angustifolia**	**S. Alberta, Canada**	**pinocembrin chalcone**	**256.07356**
1	**P. trichocarpa**	**Alaska, USA**	**2',4',6'-trihydroxydihydrochalcone**	**258.08921**
2	**P. deltoides**	**Wisconsin, USA**	**2',4',6'-trihydroxydihydrochalcone**	**258.08921**

3	**P. angustifolia**	**S. Alberta, Canada**	**2',4',6'-trihydroxydihydrochalcone**	**258.08921**
2	**P. deltoides**	**Wisconsin, USA**	**3-methyl-2-butenyl-(E)-ferulate**	**262.12051**
3	**P. angustifolia**	**S. Alberta, Canada**	**3-methyl-2-butenyl-(E)-ferulate**	**262.12051**
3	**P. angustifolia**	**S. Alberta, Canada**	**3-methyl-3-butenyl-(E)-isoferulate**	**262.12051**
3	**P. angustifolia**	**S. Alberta, Canada**	**2-methyl-2-butenyl-(E)-isoferulate**	**262.12051**
1	**P. trichocarpa**	**Alaska, USA**	**(E)-cinnamyl-(E)-cinnamate**	**264.11503**
4	**P. balsamifera**	**Białystok, Poland**	**(E)-cinnamyl-(E)-cinnamate**	**264.11503**
1	**P. trichocarpa**	**Alaska, USA**	**(E)-4-coumaryl-(E)-cinnamate**	**266.0943**
2	**P. fremontii**	**Arizona, USA**	**cinnamyl-(Z)-4-coumarate**	**266.0943**
3	**P. angustifolia**	**S. Alberta, Canada**	**tectochrysin**	**268.07356**
1	**P. trichocarpa**	**Alaska, USA**	**1-phenylethyl-(E)-4-coumarate**	**268.10995**
2	**P. deltoides**	**Wisconsin, USA**	**phenylethyl-(E)-4-coumarate**	**268.10995**
2	**P. fremontii**	**Arizona, USA**	**phenylethyl-(E)-4-coumarate**	**268.10995**
1	**P. trichocarpa**	**Alaska, USA**	**galangin**	**270.05283**
2	**P. deltoides**	**Wisconsin, USA**	**galangin**	**270.05283**
2	**P. fremontii**	**Arizona, USA**	**galangin**	**270.05283**

3	**P. angustifolia**	**S. Alberta, Canada**	**galangin**	**270.05283**
1	**P. trichocarpa**	**Alaska, USA**	**pinostrobin**	**270.08921**
1	**P. trichocarpa**	**Alaska, USA**	**pinostrobin chalcone**	**270.08921**
2	**P. deltoides**	**Wisconsin, USA**	**pinostrobin**	**270.08921**
2	**P. deltoides**	**Wisconsin, USA**	**pinostrobin chalcone**	**270.08921**
2	**P. fremontii**	**Arizona, USA**	**alpinetin**	**270.08921**
2	**P. fremontii**	**Arizona, USA**	**alpinetin chalcone**	**270.08921**
3	**P. angustifolia**	**S. Alberta, Canada**	**pinostrobin**	**270.08921**
3	**P. angustifolia**	**S. Alberta, Canada**	**pinostrobin chalcone**	**270.08921**
3	**P. angustifolia**	**S. Alberta, Canada**	**benzyl-(E)-caffeate**	**270.08921**
1	**P. fremontii**	**Arizona, USA**	**2',4',6'-α-tetrahydroxychalcone**	**272.06848**
1	**P. trichocarpa**	**Alaska, USA**	**pinobanksin**	**272.06848**
1	**P. fremontii**	**Arizona, USA**	**pinobanksin**	**272.06848**
2	**P. deltoides**	**Wisconsin, USA**	**2',4',6'-α-tetrahydroxychalcone**	**272.06848**
2	**P. deltoides**	**Wisconsin, USA**	**pinobanksin**	**272.06848**
2	**P. fremontii**	**Arizona, USA**	**naringenin**	**272.06848**

2	**P. fremontii**	**Arizona, USA**	**naringenin chalcone**	**272.06848**
3	**P. angustifolia**	**S. Alberta, Canada**	**naringenin**	**272.06848**
3	**P. angustifolia**	**S. Alberta, Canada**	**naringenin chalcone**	**272.06848**
3	**P. angustifolia**	**S. Alberta, Canada**	**naringenin chalcone**	**272.06848**
1	**P. trichocarpa**	**Alaska, USA**	**2',6'-dihydroxy-4'-methoxydihydrochalcone**	**272.10486**
2	**P. deltoides**	**Wisconsin, USA**	**2'-6'-dihydroxy-4'-methoxydihydrochalcone**	**272.10486**
3	**P. angustifolia**	**S. Alberta, Canada**	**2',6'-dihydroxy-4'-methoxydihydrochalcone**	**272.10486**
1	**P. trichocarpa**	**Alaska, USA**	**2',4',6',4-tetrahydroxydihydrochalcone**	**274.08413**
1	**P. fremontii**	**Arizona, USA**	**cinnamyl-(E)-4-coumarate**	**280.10995**
2	**P. deltoides**	**Wisconsin, USA**	**cinnamyl-(E)-4-coumarate**	**280.10995**
1	**P. fremontii**	**Arizona, USA**	**5,7-dihydroxy-3-methoxyflavanone**	**284.06848**
2	**P. deltoides**	**Wisconsin, USA**	**5,7-dihydroxy-3-methoxyflavanone**	**284.06848**
3	**P. angustifolia**	**S. Alberta, Canada**	**5,7-dihydroxy-3-methoxyflavanone**	**284.06848**
2	**P. deltoides**	**Wisconsin, USA**	**phenylethyl-(E)-caffeate**	**284.10486**
3	**P. angustifolia**	**S. Alberta, Canada**	**phenylethyl-(E)-caffeate**	**284.10486**
3	**P. angustifolia**	**S. Alberta, Canada**	**kaempferol**	**286.04774**

3	**P. angustifolia**	**S. Alberta, Canada**	**luteolin**	**286.04774**
1	**P. trichocarpa**	**Alaska, USA**	**isosakuranetin chalcone**	**286.08413**
1	**P. trichocarpa**	**Alaska, USA**	**sakauranetin**	**286.08413**
1	**P. trichocarpa**	**Alaska, USA**	**sakauranetin chalcone**	**286.08413**
2	**P. deltoides**	**Wisconsin, USA**	**isosakuranetin**	**286.08413**
2	**P. deltoides**	**Wisconsin, USA**	**isosakuranetin chalcone**	**286.08413**
3	**P. angustifolia**	**S. Alberta, Canada**	**sakauranetin**	**286.08413**
3	**P. angustifolia**	**S. Alberta, Canada**	**sakauranetin chalcone**	**286.08413**
1	**P. trichocarpa**	**Alaska, USA**	**2',6',4-trihydroxy-4'-methoxydihydrochalcone**	**288.09978**
3	**P. angustifolia**	**S. Alberta, Canada**	**catechin**	**290.07904**
4	**P. balsamifera**	**Białystok, Poland**	**n-heneicosane**	**296.3443**
4	**P. nigra**	**Białystok, Poland**	**n-heneicosane**	**296.3443**
2	**P. deltoides**	**Wisconsin, USA**	**(E)-cinnamyl-(E)caffeate**	**297.11269**
3	**P. angustifolia**	**S. Alberta, Canada**	**(E)-cinnamyl-(E)caffeate**	**297.11269**
3	**P. angustifolia**	**S. Alberta, Canada**	**phenylethyl-(iso)-ferulate**	**298.12051**
2	**P. fremontii**	**Arizona, USA**	**kaempferol-4'-methyl ether**	**300.06339**

3	**P. angustifolia**	**S. Alberta, Canada**	**quercetin**	**302.04265**
1	**P. trichocarpa**	**Alaska, USA**	**2',6'-dihydroxy-4',4-dimethoxydihydrochalcone**	**302.11543**
3	**P. angustifolia**	**S. Alberta, Canada**	**taxifolin**	**304.05831**
4	**P. balsamifera**	**Białystok, Poland**	**n-docosane**	**310.35995**
4	**P. nigra**	**Białystok, Poland**	**n-docosane**	**310.35995**
1	**P. fremontii**	**Arizona, USA**	**pinobanksin-3-acetate**	**314.07904**
2	**P. deltoides**	**Wisconsin, USA**	**pinobanksin-3-acetate**	**314.07904**
3	**P. angustifolia**	**S. Alberta, Canada**	**pinobanksin-3-acetate**	**314.07904**
3	**P. angustifolia**	**S. Alberta, Canada**	**pinobanksin-3-acetate chalcone**	**314.07904**
3	**P. angustifolia**	**S. Alberta, Canada**	**quercetin-7-methyl ether**	**316.05831**
4	**P. balsamifera**	**Białystok, Poland**	**n-tricosane**	**324.3756**
4	**P. nigra**	**Białystok, Poland**	**n-tricosane**	**324.3756**
2	**P. fremontii**	**Arizona, USA**	**5,7-dihydroxy-3-butanyloxyflavone**	**326.11543**
1	**P. fremontii**	**Arizona, USA**	**pinobanksin-3-propanoate**	**328.09469**
2	**P. deltoides**	**Wisconsin, USA**	**pinobanksin-3-propanoate**	**328.09469**
3	**P. angustifolia**	**S. Alberta, Canada**	**pinobanksin-3-propanoate**	**328.09469**

3	**P. angustifolia**	**S. Alberta, Canada**	**quercetin-7,3'-dimethyl ether**	**330.07396**
4	**P. balsamifera**	**Białystok, Poland**	**n-tetracosane**	**338.39125**
4	**P. nigra**	**Białystok, Poland**	**n-tetracosane**	**338.39125**
1	**P. fremontii**	**Arizona, USA**	**pinobanksin-3-butanoate**	**342.11034**
2	**P. deltoides**	**Wisconsin, USA**	**pinobanksin-3-butanoate**	**342.11034**
3	**P. angustifolia**	**S. Alberta, Canada**	**pinobanksin-3-butanoate**	**342.11034**
4	**P. balsamifera**	**Białystok, Poland**	**n-pentacosane**	**352.4069**
4	**P. nigra**	**Białystok, Poland**	**n-pentacosane**	**352.4069**
1	**P. fremontii**	**Arizona, USA**	**pinobanksin-3-pentanoate**	**356.12599**
2	**P. deltoides**	**Wisconsin, USA**	**pinobanksin-3-pentanoate**	**356.12599**
3	**P. angustifolia**	**S. Alberta, Canada**	**pinobanksin-3-pentanoate**	**356.12599**
4	**P. balsamifera**	**Białystok, Poland**	**n-hexacosane**	**366.42255**
4	**P. nigra**	**Białystok, Poland**	**n-hexacosane**	**366.42255**
2	**P. fremontii**	**Arizona, USA**	**pinobanksin-3-hexanoate**	**370.1416**
4	**P. balsamifera**	**Białystok, Poland**	**n-heptacosane**	**380.4382**
4	**P. nigra**	**Białystok, Poland**	**n-heptacosane**	**380.4382**

4	**P. balsamifera**	**Białystok, Poland**	**n-octacosane**	**394.45385**
4	**P. nigra**	**Białystok, Poland**	**n-octacosane**	**394.45385**
4	**P. balsamifera**	**Białystok, Poland**	**n-nonacosane**	**408.4695**
4	**P. nigra**	**Białystok, Poland**	**n-nonacosane**	**408.4695**
4	**P. balsamifera**	**Białystok, Poland**	**n-triacontane**	**422.48515**
4	**P. nigra**	**Białystok, Poland**	**n-triacontane**	**422.48515**
4	**P. balsamifera**	**Białystok, Poland**	**n-hentriacontane**	**436.5008**
4	**P. nigra**	**Białystok, Poland**	**n-hentriacontane**	**436.5008**
2	**P. fremontii**	**Arizona, USA**	**2',4',6',α-tetrahydroxychalcone**	**---**
3	**P. angustifolia**	**S. Alberta, Canada**	**diprenyl caffeate**	**---**
3	**P. angustifolia**	**S. Alberta, Canada**	**2',6',α-trihydroxy-4'-methoxychalcone**	**---**
3	**P. angustifolia**	**S. Alberta, Canada**	**3-methylbutyl-(E)-caffeate**	**---**
3	**P. angustifolia**	**S. Alberta, Canada**	**2-methylproply-(E)-caffeate**	**---**

Appendix C

R script for comparative metabolomics data analysis

This script is a conservative, open source data analysis procedure for metabolomics experiments that use a ‘composite’ quality control sample, as described in the Chapter 2 results subsection “Honey bees collect resin from balsam poplar and eastern cottonwood.”

```
#############################################################
## Setting up filepaths, names, and packages
## Before you start - Your active directory should contain all your data files and a
## directory called 'results'
## You will need to download the following packages: ‘xcms’, ‘RANN’, and ‘lattice’

##Load dependent packages
library(xcms)
library(RANN)
library(lattice)

## about.csv is a file with information about the raw data files in the 'results' directory
## about.csv needs 4 columns: id (a unique identifier), group (sample treatment, ect),
Nfile (neg #mode file path), Pfile (pos mode file path)
## Your quality control group must be labeled as “Composite” in the group column

## Set up a name and read in your 'about' file
file.about <- "results/about.csv"
ab <- read.csv(file.about, as.is=TRUE)

## Set the directory where XCMS will look for your data files
## The filepath for this directory should be given between the quotation marks
dir.raw <- "/project/hegemana/Mike Wilson/Species analysis" ab$Nfile <-
file.path(dir.raw, ab$Nfile)
ab$Pfile <- file.path(dir.raw, ab$Pfile)

## Set the directory to store script output (your 'results' directory) and set up other
filenames
dir.results <- "results"
file.peaksP <- file.path(dir.results, "peaksP.Rdata")
file.peaksN <- file.path(dir.results, "peaksN.Rdata")
file.groupsP <- file.path(dir.results, "groupsP.Rdata")
file.groupsN <- file.path(dir.results, "groupsN.Rdata")
```

```
file.intb <- file.path(dir.results, "intb.csv")
file.Rsq <- file.path(dir.results, "Rsq.csv")
file.scree <- file.path(dir.results, "scree.pdf")
file.screecsv <- file.path(dir.results, "scree.csv")
file.para <- file.path(dir.results, "para.pdf")
file.pc12 <- file.path(dir.results, "pc12.pdf")
file.pc12_modified <- file.path(dir.results, "pc12_modified.pdf")

## Make sure all the files in about.csv exist
## If a file does not 'exist' there is probably a discrepancy between the file name in
about.csv vs. #the actual file name
file.exists(ab$Nfile) #Make sure all the (-) files exist
file.exists(ab$Pfile) #Make sure all the (+) files exist

#############################################################
## Have XCMS get peaks from raw MS data and group them.
## Do not include files that are all noise (e.g. blanks), as the analysis will not work
## These steps can take time (depending on your computing power), so results are saved
## as R data files.  This section can be skipped later if you want to change the
## downstream analysis.

## Get positive peaks from raw MS data and group them
xset <- xcmsSet(ab$Pfile, method="centWave", ppm=10, peakwidth=c(5,50),
            fitgauss=TRUE, verbose.columns=TRUE)
save(xset, file=file.peaksP)
xset <- group(xset, method="nearest", mzCheck=2, rtCheck=5)
save(xset, file=file.groupsP)

## Get negative peaks  from raw MS data and group them
xset <- xcmsSet(ab$Nfile, method="centWave", ppm=10, peakwidth=c(5,50),
            fitgauss=TRUE, verbose.columns=TRUE)
save(xset, file=file.peaksN)
xset <- group(xset, method="nearest", mzCheck=2, rtCheck=5)
save(xset, file=file.groupsN)

#############################################################
## Merge (+) and (-) peak lists and create the final peak matrix

## Get positive peak list
load(file.groupsP)
Pintb <- groupval(xset, value="intb")
Pintb <- Pintb[,match(ab$Pfile, filepaths(xset))]
rownames(Pintb) <- paste("P", groupnames(xset), sep=".")
```

```
colnames(Pintb) <- ab$id
dim(Pintb)

## Get negative peak list
load(file.groupsN)
Nintb <- groupval(xset, value="intb")
Nintb <- Nintb[,match(ab$Nfile, filepaths(xset))]
rownames(Nintb) <- paste("N", groupnames(xset), sep=".")
colnames(Nintb) <- ab$id
dim(Nintb)

## Merge lists together
intb <- rbind(Pintb, Nintb)
dim(intb)

## Count the number of peaks in each raw data file and remove files with less than 100
## peaks from the analysis.  Using files with < 100 peaks prevents completion of the
## analysis.
ab$npeaks <- colSums(!is.na(intb))
intb <- intb[,colnames(intb) %in% ab$id[ab$npeaks >= 100]]
dim(intb)

## Remove peaks that don't appear in every composite sample.  This is a conservative
## data filter that is designed to exclude noise peaks from the analysis.
comp <- intb[, ab$id[ab$group=="Composite"], drop=FALSE]
ncomp <- rowSums(!is.na(comp))
intb <- intb[ncomp >= ncol(comp), , drop=FALSE]
dim(intb)

## Remove composite samples.  They are not needed for the rest of the analysis.
intb <- intb[, !colnames(intb) %in% ab$id[ab$group=="Composite"], drop=FALSE]
dim(intb)

## Remove composite peaks that are always or never present in the whole data set.  We
## only want to look at things that are different between sample groups.
bsum <- rowSums(is.na(intb))
intb <- intb[bsum>0 & bsum<ncol(intb), , drop=FALSE]
dim(intb)

## Make all missing values in the peak matrix = 0 (e.g. if a peak was not present in a file,
## it will get an area value of 0)
intb[is.na(intb)] <- 0
```

```
##Create 'intb.csv'.  This is your peak matrix used for analysis
write.csv(intb, file=file.intb, na="")

#################################################################
## Modify 'about.csv' so that it matches 'intb.csv'.  This is important for tracking the
## meta-data in ‘about.csv’ correctly
intb <- as.matrix(read.csv(file.intb, row.names=1, check.names=FALSE))
intb.saved <- intb

## Remove composite files and files with < 100 peaks from "about.csv"
ab2 <- ab[match(colnames(intb), ab$id),]
ab2$group <- factor(ab2$group)
ab2.saved <- ab2

#################################################################
## Optional - Using a subset of your peak matrix for analysis instead of the whole thing.
## This allows you to change the files you want to compare (e.g. eliminating some
## groups from the analysis) without having to create a new ‘about.csv. and run your files
## through XCMS a second #time.  Only use one of the options below – isolating single
## groups or isolating multiple groups

## How to isolate a single group in your peak matrix, ‘intb’, for analysis.
## The group (as given in ‘about.csv’) you want to analyze should be in the quotation
## marks
intb <- intb.saved[,grep("P. deltoides", ab2$group)]

## How to isolate multiple groups in your peak matrix, ‘intb’, for analysis.
##The groups (as given in ‘about.csv’) you want to analyze should be in the quotation
## marks
##You can add another group into the comparison using '|'
use <- grepl("P. deltoides May", ab2$group) | grepl("P. deltoides June", ab2$group) use
<- grep("P. deltoides", ab2$group)

## change both 'intb.csv' and 'about.csv' to include only those groups we want to analyze
intb <- intb.saved[,use]
ab2 <- ab2.saved[use,]

#################################################################
##Analysis of your peak matrix, ‘intb.csv’.  Only use one principle component analysis
##(PCA).

## Prepare 'intb.csv' for presence/absence principle component analysis (PCA).  This use
## of PCA is designed to highlight differences between samples due characteristic and
```

```
## exclusive peaks.
Bx <- (intb>0)*1
Bx <- Bx[rowMeans(Bx) > 0 & rowMeans(Bx)<1,]
B <- scale(t(Bx))

## Get principal components
Bs <- svd(B, nv=0)
Bpcs <- Bs$u %*% diag(Bs$d)
Bvar <- Bs$d^2/nrow(B)
colnames(Bpcs) <- paste("PC", 1:ncol(Bpcs), sep="")

## Create a scree plot showing the % variation explained by each principal component.
## Creates both a .pdf and .csv file of scree plot results.
npc <- 30 # keep this many PCs for evaluation, could keep more or less
evar <- Bvar/sum(Bvar)
pdf(file.scree) #creates the scree plot as a pdf file in the 'results' directory
plot(evar[1:npc], ylim=c(0, max(evar)*1.05), yaxs="i")
dev.off()
write.csv(data.frame(pc=1:npc, var=evar[1:npc]), row.names=FALSE, file=file.screecsv)

## Create a .pdf plot of the first two principal components.  You can change the principle
## components plotted by modifying 'PC2~PC1' to your desired principle components
p <- xyplot(PC2~PC1, group=ab2$group, data=data.frame(Bpcs),
auto.key=list(space="right"))
pdf(file.pc12)
plot(p)
dev.off()

## If there is an outlier in your PCA plot, determine which sample is it.  Your definition
## of 'outlier' should be governed by your PCA plot, but for this example we have
## chosen a distance magnitude of 30 to define 'outlier'.
ab2[which(Bpcs[,"PC1"]>30),]

## Determine the row number of any outliers in 'about.csv'
which(Bpcs[,"PC1"]>30)

## Create a new PCA plot that does not include the outlier(s).  19 is the outlier row
## number in 'about.csv' for this example.
p <- xyplot(PC2~PC1, group=ab2$group[-19], data=data.frame(Bpcs[-19,]),
auto.key=list(space="right"))
pdf(file.pc12_modified)
plot(p)
dev.off()
```

```
## Create a parallel plot of first 5 principle components.  You can exclude outliers by
## modifying ab2$group to ab2$group[-"outlier row"]
p <- parallelplot(~Bpcs[,1:5], group=ab2$group[-19], horizontal=FALSE,
common.scale=TRUE, auto.key=list(space="right")) pdf(file.para)
plot(p)
dev.off()

## Determine if any peaks are strongly correlated with the principle components, and get
## the R^2 values for this correlation.  This will essentially tell you which peaks are
## causing any group segregation seen in the PCA plot.  The default is to use only the
## first 3 principle components.
npc2 <- 3
pcs <- Bpcs[,1:npc2]
sumpcs <- colSums(pcs^2)
Rsq <- colSums(cor(pcs, B)^2*sumpcs)/sum(sumpcs)

## Determine the presence/absence of a given peak by sample.  This is designed to give
## you an overview of peaks that are characteristic of/reproducible in a given sample
## group.
## Outliers can be excluded by modifying ab2$group to ab2$group[-"outlier row"]
ii <- split(1:ncol(Bx), ab2$group)
count <- sapply(ii, function(i) rowSums(Bx[,i,drop=FALSE]))

## Output peak correlation and sample appearance information (ordered by highest R^2
## value) into a .csv file.  This gives you a file that can be manipulated in Microsoft
## Excel to find reproducible and characteristic peaks that are contributing to group
## segregation in the PCA plot.
out <- cbind(count, Rsq)
out <- out[order(out["Rsq"],decreasing=TRUE),]
write.csv(out, file=file.Rsq)
```